

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 12, 2005, 07:49:31 ; Search time 39.2 Seconds
(without alignments)
69.064 Million cell updates/sec

Title: US-10-691-330-1

Perfect score: 42

Sequence: 1 MQPPPLP 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	42	100.0	7	4	AAB72500 Colostrin
2	42	100.0	7	4	AAB59324 Ewe colos
3	42	100.0	7	4	AAB72246 Colostrin
4	42	100.0	7	4	AAB72532 Colostrin
5	42	100.0	7	5	AAO14577 Neural ce
6	42	100.0	7	5	AAM51036 Colostrin
7	42	100.0	7	5	AAE20228 Colostrin
8	42	100.0	7	8	ADN60295 Constitue
9	42	100.0	7	8	ADS74400 Ovine col
10	42	100.0	8	4	AAB59354 Ewe colos
11	42	100.0	316	4	ABB67510 Drosophil
12	40	95.2	785	8	ADN72669 Thale cre
13	39	92.9	9	8	ADK10336 Human pap
14	39	92.9	10	8	ADK10337 Human pap
15	39	92.9	10	8	ADK10338 Human pap
16	39	92.9	95	4	AAU45697 Propionib
17	39	92.9	95	6	ABM42216 Propionib
18	39	92.9	126	5	ADK34767 Novel hum
19	39	92.9	132	4	AAM17529 Peptide #
20	39	92.9	132	4	ABB36552 Peptide #
21	39	92.9	132	4	AAM30051 Peptide #
22	39	92.9	132	4	ABB31347 Peptide #
23	39	92.9	132	4	ABB21889 Protein #
24	39	92.9	132	4	AAM69716 Human bon
25	39	92.9	132	4	AAM57317 Human bra

26	39	92.9	132	4	ABG51400 Human liv
27	39	92.9	132	4	AAM05201 Peptide #
28	39	92.9	132	5	ABG39337 Human pep
29	39	92.9	572	8	ADN26842 Bacterial
30	39	92.9	572	8	ADN26614 Bacterial
31	39	92.9	612	8	ADS28506 Bacterial
32	38	90.5	256	7	ADJ71090 Human hea
33	38	90.5	256	8	ADS95110 Human the
34	38	90.5	515	5	ABB90945 Herbicida
35	38	90.5	662	8	ABM81932 Tumour-as
36	37	88.1	8	8	ADK10335 Human pap
37	37	88.1	38	4	AAE10123 Human ion
38	37	88.1	38	5	AAU83499 Novel hum
39	37	88.1	38	6	ABU97384 Amino aci
40	37	88.1	38	7	ADE29306 Novel hum
41	37	88.1	70	5	ABP10528 Human ORF
42	37	88.1	73	3	AAB41170 Human ORF
43	37	88.1	73	5	ABP01103 Human ORF
44	37	88.1	95	4	ABG21168 Novel hum
45	37	88.1	99	3	AAG23978 Arabidops

ALIGNMENTS

RESULT 1
AAB72500
ID AAB72500 standard; peptide; 7 AA.
XX AAB72500;
XX AC
XX 09-MAY-2001 (first entry)
XX Colostrinin peptide #1.
XX Colostrinin peptide #1.
XX Dermatological; oxidative stress regulator; colostrinin.
XX Unidentified.
XX OS
XX W0200112650-A2.
XX
XX 22-FEB-2001.
XX
XX 17-AUG-2000; 2000WO-US022665.
XX
XX 17-AUG-1999; 99US-0149310P.
XX (TEXA) UNIV TEXAS SYSTEM.
XX Stanton GJ, Hughes TK, Boldogh I;
XX WPI; 2001-218342/22.
XX
XX Modulating oxidative stress level in a cell, involves contacting the cell with an oxidative stress regulator selected from colostrinin, its constituent peptide, analog or their combinations.
XX
XX Claim 6; Page 25; 48pp; English.
XX
XX The present invention relates to a method for modulating the oxidative stress level in a cell or a patient, comprising contacting the cell with, or administering to the patient, an oxidative stress regulator selected from colostrinin, or its constituent peptide (e.g. the present peptide), to change the level of an oxidising species in the cell. The method can be used to treat oxidative damage to skin, by decreasing or preventing an increase in the level of damage to a biomolecule of the patient

XX Sequence 7 AA;

Query Match 100.0%; Score 42; DB 4; Length 7;
Best Local Similarity 100.0%; Pred. NO. 1.8e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

PA (TEXA) UNIV TEXAS SYSTEM.
XX Boldogh I;
XX WPI; 2001-226545/23.
XX Use of colostrin, its constituent peptide or analog as a neural cell
PT regulator, for promoting neural cell differentiation and treating damaged
PT neural cells in a patient.
XX Claim 6; Page 21; 35pp; English.
PS The present invention relates to a method for promoting neural cell
XX differentiation and treating damaged neural cells, using colostrin and
CC colostrin constituent peptides (e.g. the present peptide) as a neural
CC cell regulator. Colostrin is a polypeptide complex found in colostrum
XX cell regulator. Colostrin is a polypeptide complex found in colostrum
SQ Sequence 7 AA;
Query Match 100.0%; Score 42; DB 4; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.8e+06; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MQPPPLP 7
Db 1 MQPPPLP 7
RESULT 5
AAO14577
ID AAO14577 standard; peptide; 7 AA.
AC AAO14577;
XX 27-MAY-2002 (first entry)
DT Neural cell regulatory colostrin peptide 1.
DE Neural cell differentiation; neural cell regulator; colostrin peptide;
KW neural cell formation; proline-rich polypeptide aggregate; colostrum;
XW neural cell treatment.
XX Unidentified.
OS Key Location/Qualifiers
FH Modified-site 7 /note= "Optional C-terminal amide"
FT WO200213851-A1.
XX 21-FEB-2002.
XX 17-AUG-2000; 2000WO-US022777.
XX 17-AUG-2000; 2000WO-US022777.
XX (TEXA) UNIV TEXAS SYSTEM.
PA Boldogh I, Stanton JG, Hughes TK;
XX WPI; 2002-269152/31.
XX Promoting cell differentiation in a patient involves use of blood cell
PT regulator selected from colostrin, its constituent peptide and/or
PT analog.
XX Claim 7; Page 21; 37pp; English.
PS The invention comprises a method for promoting cell differentiation (e.g.
XX neural cell differentiation). The method involves contacting cells with a
CC neural cell regulator (i.e. a colostrin peptide) in order to change the
CC cells in morphology to form neural cells. Colostrin is a proline-rich
CC polypeptide aggregate that is present in colostrum. The method of the

CC invention is useful for promoting the differentiation of cells and for
CC treating damaged neural cells in a patient. The present amino acid
CC sequence represents a specifically claimed colostrin peptide used in
CC the method of the invention
XX Sequence 7 AA;
Query Match 100.0%; Score 42; DB 5; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.8e+06; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MQPPPLP 7
Db 1 MQPPPLP 7
RESULT 6
AAM51036
ID AAM51036 standard; peptide; 7 AA.
XX AAM51036;
AC AAM51036;
XX 30-MAY-2002 (first entry)
DT Colostrin constituent peptide.
DE Colostrin; colostrum; immunomodulator; cardiovascular;
XW blood cell regulator; cytokine inducer; beta-casein; human.
XX Homo sapiens.
OS Key Location/Qualifiers
FH Modified-site 7 /note= "optional C-terminal amidation"
FT WO200213849-A1.
XX 21-FEB-2002.
XX 17-AUG-2000; 2000WO-US022775.
XX 17-AUG-2000; 2000WO-US022775.
XX (TEXA) UNIV TEXAS SYSTEM.
PA (REG-) REGEN THERAPEUTICS PLC.
XX Stanton GJ, Hughes TK, Boldogh I, Georgiades J;
XX WPI; 2002-269150/31.
XX Modulation of blood cell proliferation in a patient involves use of blood
PT cell regulator selected from colostrin, its constituent peptide and/or
PT analog.
XX Claim 1; Page 34; 54pp; English.
PS The present sequence is that of a colostrin constituent peptide that is
CC preferred for use as an immunological regulator and as a blood cell
CC regulator in claimed methods of the invention. It is classified as having
CC a beta-casein homologue precursor. Methods are claimed for: inducing a
CC cytokine in a cell by contact with an immunological regulator, where the
CC cell is present in a cell culture, a tissue, an organ or an organism, and
CC the cell is mammalian, including human; modulating an immune response in
CC a cell by contact with the immunological regulator under conditions
CC effective to induce a cytokine; modulating an immune response in a
CC patient by administering an immunological regulator under conditions
CC effective to induce a cytokine, where the immunological regulator is
CC administered topically or as part of a dietary supplement, and where the
CC immune response is specific or non specific, an interferon response or an
CC antibody response; modulating blood cell proliferation by contacting
CC blood cells with a blood cell regulator, where the blood cells are
CC present in a cell culture or an organism, are mammalian or human, and
CC where the blood cells are increased in number or differentiated; and a

CC method for modulating blood cell proliferation in a patent. A claimed
 CC cytokine-inducing composition comprises a pharmaceutical carrier and an
 CC active agent such as the present peptide. Cytokines induced by this
 CC peptide in human leucocyte cultures include interferon-gamma, tumour
 CC necrosis factor-alpha, interleukin-4, interleukin-6, interleukin-10 and
 CC interleukin-12. It was one of the best overall inducers in almost all
 CC cytokine and blood cell proliferation experiments conducted

XX Sequence 7 AA;
 SQ Query Match 100.0%; Score 42; DB 5; Length 7;
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MQPPPLP 7
 Db 1 MQPPPLP 7

RESULT 7
 AAE20228
 ID AAE20228 standard; peptide; 7 AA.
 XX AC AAE20228;
 DT 18-JUN-2002 (first entry)
 XX DE Colostrinin constituent peptide #1.
 XX KW Blood cell regulator; colostrinin; constituent peptide; oxidative stress;
 KW therapy; oxidative damage; skin; aging; wound healing; cell replacement;
 KW tissue; organ; cosmetic procedure; repair; regeneration; preservation;
 KW transplantation; implantation; dermatological; vulnary.

XX Unidentified.
 OS Key Location/Qualifiers
 XX Modified-site 7 /note= "Optionally C-terminal amide"
 FT FT
 XX WO200213850-A1.
 XX 21-FEB-2002.
 XX 17-AUG-2000; 2000WO-US022776.
 XX 17-AUG-2000; 2000WO-US022776.
 XX (TEXA) UNIV TEXAS SYSTEM.
 XX Stanton GJ, Hughes TK, Boldogh I;
 XX WPI; 2002-269151/31.
 XX Composition useful for the modulation of blood cell proliferation in a
 PT patient comprises a blood cell regulator selected from colostrinin, its
 PT constituent peptide and/or analog.
 PS Claim 6; Page 25; 51pp; English.

XX The invention relates to a composition which comprises a blood cell
 CC regulator selected from colostrinin, its constituent peptide and/or
 CC analogue. The invention is used for modulating the oxidative stress level
 CC in a cell e.g. mammalian or human cell present in a cell culture, tissue,
 CC organ, or organism; or for treating oxidative damage to the skin of a
 CC patient e.g. animal or human; to modulate oxidative stress during/ after
 CC a premature birth or normal birth, preventing/delaying aging in a
 CC patient, enhancing wound healing, and the reduction of side effects of
 CC cosmetic procedures. The method changes the level of an oxidizing species
 CC in the cell, such as decreases or prevents increase in the level of
 CC damage to a biomolecule of the patient selected from DNA, protein and/or
 CC lipid, compared to the same conditions when the oxidative stress
 CC regulator is not present. The modulation of oxidative stress results in

CC enhanced repair, regeneration, and replacement of cells, tissues and
 CC organs (e.g. kidney, liver, pancreas, skin, and the other internal and
 CC external organs), as well as enhanced preservation of such organs for
 CC transplantation, implantation, or scientific research. The present
 CC sequence is a colostrinin constituent peptide
 XX
 SQ Sequence 7 AA;

Query Match 100.0%; Score 42; DB 5; Length 7;
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MQPPPLP 7
 Db 1 MQPPPLP 7

RESULT 8
 ADN60295
 ID ADN60295 standard; peptide; 7 AA.
 XX AC ADN60295;
 DT 29-JUL-2004 (first entry)
 XX DE Constituent peptide of colostrinin SEQ ID NO:1.
 XX KW modulator; colostrinin; intracellular signaling molecule modulator;
 KW 4HNE-mediated lipid peroxidation down regulation; apoptosis inhibition;
 KW DNA damage; beta-amyloid; retinoic acid; cytosolic; 4HNE inhibitor;
 KW 4HNE-protein adduct formation reduction;
 KW 4HNE-mediated glutathione depletion inhibition; p53 inhibition;
 KW c-Jun NH2-terminal kinase inhibition.

XX Synthetic.
 OS WO2004037851-A2.
 XX 06-MAY-2004.
 XX 22-OCT-2003; 2003WO-US033423.
 XX 22-OCT-2002; 2002US-0420369P.
 XX (TEXA) UNIV TEXAS SYSTEM.
 XX (BOLD/) BOLDOGH I.
 XX (STAN/) STANTON J G.
 XX (GEOR/) GEORGIADIS J A.
 XX (HUGH/) HUGHES T K.
 XX (KRUIZ/) KRUIZEL M.
 XX Boldogh I, Stanton JG, Georgiades JA, Hughes TK, Kruzel M;
 XX WPI; 2004-365494/34.
 XX Use of colostrinin for e.g. modulating an intracellular signaling
 PT molecule in a cell, down regulating 4HNE-mediated lipid peroxidation in a
 PT cell, inhibiting apoptosis in a cell, or protecting against DNA damage in
 PT a cell.

XX Claim 6; SEQ ID NO 1; 46pp; English.
 XX The present invention describes the use of a modulator selected from
 CC colostrinin, its constituent peptide, its active analogue, and a
 CC combination of these, for modulating an intracellular signaling molecule
 CC in a cell, down regulating 4HNE-mediated lipid peroxidation in a cell,
 CC inhibiting apoptosis in a cell, protecting against DNA damage in a cell,
 CC and reducing the toxic effect of beta-amyloid or retinoic acid on a cell.
 CC The modulator has cytosolic activity, and can be used as a 4HNE
 CC inhibitor. The modulator is useful in the manufacture of a medicament for
 CC reducing 4HNE-protein adduct formation, inhibiting 4HNE-mediated
 CC glutathione depletion, inhibiting 4HNE-induced activation of p53 protein,
 CC and/or inhibiting 4HNE-induced activation of c-Jun NH2-terminal kinases.

CC Colostrin, or its constituent peptide or active analogue is useful for
 CC inhibiting apoptosis in a cell, protecting against DNA damage in a cell,
 CC or reducing the toxic effect of beta-amyloid or retinoic acid on a cell.
 CC The present sequence represents a synthetic constituent peptide of
 CC colostrin, which can be used as a modulator in the present invention.
 CC
 XX
 SQ Sequence 7 AA;

Query Match 100.0%; Score 42; DB 8; Length 7;
 Best Local Similarity 100.0%; Pred. No. 1.8e+06; Indels 0; Gaps 0;
 Matches 7; Conservative 0; Mismatches 0;

Qy 1 MQPPPLP 7
 |||||
 Db 1 MQPPPLP 7

RESULT 9

AD574400
 ID AD574400 standard; peptide; 7 AA.

XX
 AC AD574400;

XX
 DT 16-DEC-2004 (first entry)

XX
 DE Ovine colostrin peptide.

XX
 KW Colostrum; colostrin; sheep; peptide purification.

XX
 OS Ovis aries.

XX
 PN WO2004081038-A1.

XX
 PD 23-SEP-2004.

XX
 PF 10-MAR-2004; 2004WO-GB001014.

XX
 PR 11-MAR-2003; 2003GB-00005552.

XX
 PR 08-MAR-2004; 2004GB-00005190.

XX
 PA (REGG-) REGEN THERAPEUTICS PLC.

XX
 PI Georgiades JA, Polanowski A, Wilusz T, Kruzel ML;

XX
 DR WPI; 2004-677519/66.

XX
 PT Recovering peptides such as colostrin from mammalian colostrum, by
 PT mixing colostrum with alcohol to form alcohol phase containing peptides
 PT and precipitate, separating alcohol phase from precipitate, and
 PT recovering alcohol phase.

XX
 PS Disclosure; SEQ ID NO 17; 41pp; English.

XX
 CC The present sequence is that of a peptide that can be recovered from
 CC ovine colostrin using the method of the invention. The invention
 CC provides a method for the recovery of peptides (especially colostrin)
 CC from colostrum in substantially pure, biologically active form and in
 CC high yield. The method involves mixing the colostrum with an alcohol to
 CC form an alcohol phase containing the colostrin and a precipitate
 CC containing higher molecular weight caseins and other proteins. Best
 CC results are obtained using methanol or ethanol of at least 80%, and
 CC preferably up to 100% purity. The alcohol phase is then separated from
 CC the precipitate, and the colostrin is separated from the alcohol.
 CC preferably by evaporation, to form a colostrin-rich phase, which is
 CC recovered. A precipitation agent, such as ammonium sulfate, may be added
 CC either to the alcohol phase or, preferably, to the colostrin-rich phase
 CC to induce precipitation of the colostrin peptides. The method is
 CC generally applicable to the separation of peptides from fluids containing
 CC higher molecular weight proteins, lipids, carbohydrates and/or nucleic
 CC acids.

XX
 SQ Sequence 7 AA;

Query Match 100.0%; Score 42; DB 8; Length 7;
 Best Local Similarity 100.0%; Pred. No. 1.8e+06; Indels 0; Gaps 0;
 Matches 7; Conservative 0; Mismatches 0;

Qy 1 MQPPPLP 7
 |||||
 Db 1 MQPPPLP 7

RESULT 10

AAB59354
 ID AAB59354 standard; peptide; 8 AA.

XX
 AC AAB59354;

XX
 DT 21-MAR-2001 (first entry)

XX
 DE Ewe colostrin peptide fragment derived sequence #14.

XX
 KW Sheep; colostrin; proline rich polypeptide; colostrum; immune disorder;
 KW central nervous system disorder; dietary supplement; beta-amyloid plaque.

XX
 OS Ovis sp.

XX
 PN WO200075173-A2.

XX
 PD 14-DEC-2000.

XX
 PF 02-JUN-2000; 2000WO-GB002128.

XX
 PR 02-JUN-1999; 99GB-00012852.

XX
 PA (REGG-) REGEN THERAPEUTICS PLC.

XX
 PI Georgiades JA;

XX
 DR WPI; 2001-071058/08.

XX
 PT Peptides having an N-terminal amino acid sequence isolated from
 PT colostrin for treating e.g. disorders of the central nervous system and
 PT immune system, viral and bacterial infections, and diseases characterized
 PT by amyloid plaques.

XX
 PS Claim 8; Page 27; 63pp; English.

XX
 CC The present invention provides the sequences of a number of peptides
 CC found in ewe's colostrin. Colostrin is the proline-rich polypeptide
 CC fragment of colostrum. These peptides can be used in the treatment of
 CC central nervous system disorders such as senile dementia, Parkinson's
 CC disease, Alzheimer's disease, psychosis and neurosis, immune system
 CC disorders such as bacterial and viral infections, to improve the
 CC development of a child's immune system, as a dietary supplement, and to
 CC promote the dissolution of beta-amyloid plaques

XX
 SQ Sequence 8 AA;

Query Match 100.0%; Score 42; DB 4; Length 8;
 Best Local Similarity 100.0%; Pred. No. 1.8e+06; Indels 0; Gaps 0;
 Matches 7; Conservative 0; Mismatches 0;

Qy 1 MQPPPLP 7
 |||||
 Db 2 MQPPPLP 8

RESULT 11

ABB67510

ID ABB67510 standard; protein; 316 AA.

XX
 AC ABB67510;

XX
 DT 26-MAR-2002 (first entry)

XX

DE Drosophila melanogaster polypeptide SEQ ID NO 29322.
 XX Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 XX Drosophila melanogaster.
 XX WO200171042-A2.
 XX 27-SEP-2001.
 XX 23-MAR-2001; 2001WO-US009231.
 XX 23-MAR-2000; 2000US-0191637P.
 PR 11-JUL-2000; 2000US-00614150.
 XX (PEKE) PE CORP NY.
 PA Venter JC, Adams M, Li PWD, Myers EW;
 XX WPI; 2001-656860/75.
 DR N-PSDB; ABL11613.
 XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions.
 XX Disclosure; SEQ ID NO 29322; 21pp + Sequence Listing; English.
 XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins (AB557737-
 CC AB572072). The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 316 AA;
 Query Match 100.0%; Score 42; DB 4; Length 316;
 Best Local Similarity 100.0%; Pred. No. 3.4e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MQPPPLP 7
 DB 249 MQPPPLP 255
 RESULT 12
 ADN72669
 ID ADN72669 standard; protein; 785 AA.
 XX AC ADN72669;
 XX 15-JUL-2004 (first entry)
 DT Thale cress protein upregulated in E2Fa/Dpa expressing plants SeqID 564.
 DE plant; transgenic; E2Fa/Dpa transcription factor; growth regulator;
 KW animal feed product; thale cress; cell wall biosynthesis;
 KW nitrogen metabolism; carbon metabolism.
 XX Arabidopsis thaliana.
 XX WO2004035798-A2.
 PN 29-APR-2004.
 PD 20-OCT-2003; 2003WO-EP011658.
 XX

PR 18-OCT-2002; 2002EP-00079408.
 XX (CROP-) CROPDESIGN NV.
 XX Inze D, De Veylder L, Vlieghe K;
 PI WPI; 2004-348466/32.
 DR N-PSDB; ADN72668.
 XX Altering plant characteristics, useful for producing plants for enzyme or
 PT pharmaceutical production comprises modifying in a plant, expression of
 PT one or more nucleic acids and/or modifying level or activity of one or
 XX more proteins.
 XX Claim 1; SEQ ID NO 564; 134pp; English.
 XX This invention relates to a novel method for altering one or more plant
 CC characteristics. Specifically, it refers to identifying genes that are up
 CC - or down-regulated in transgenic plants overexpressing the heterodimeric
 CC E2Fa/Dpa transcription factor of Arabidopsis and using these sequences to
 CC alter plant characteristics accordingly. The present invention describes
 CC generating transgenic plants for the production of growth regulators,
 CC enzymes, therapeutics, pharmaceuticals and animal feed products, where
 CC the altered plant characteristics are selected from increased yield or
 CC biomass, enhanced survival capacity, stress tolerance, plant architecture
 CC or physiology, altered endoreplication, biochemistry, signal
 CC transduction, storage lipid mobilisation and/or altered photosynthesis,
 CC each relative to the corresponding wild type plants. Accordingly, these
 CC sequences can also be useful as positive or negative selectable markers
 CC during transformation of cells or tissues. The identified genes play a
 CC role in a variety of biological processes such as DNA replication, cell
 CC wall biosynthesis, nitrogen and/or carbon metabolism or they function as
 CC transcription factors. This polypeptide sequence is thale cress protein
 CC expressed by a gene upregulated 1.3 fold or more in plants overexpressing
 CC the E2Fa/Dpa transcription factor, given in an exemplification of the
 XX invention.
 SQ Sequence 785 AA;
 Query Match 95.2%; Score 40; DB 8; Length 785;
 Best Local Similarity 85.7%; Pred. No. 1.4e+03;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MQPPPLP 7
 DB 652 MQPPPLP 658
 RESULT 13
 ADK10336
 ID ADK10336 standard; peptide; 9 AA.
 XX AC ADK10336;
 XX 06-MAY-2004 (first entry)
 DT Human papillomavirus peptide #2391.
 DE pathogenic virus; alternative reading frame; antigenic determinant;
 DE virucide; vaccine; therapeutic agent; infection; HPV.
 KW Human papillomavirus.
 OS WO2004011650-A2.
 PN 05-FEB-2004.
 PD 24-JUL-2003; 2003WO-EP008112.
 XX 24-JUL-2002; 2002AT-00001124.
 PR 11-JUL-2003; 2003EP-00450171.
 XX (INTE-) INTERCELL AG.
 PA

XX Mattner F, Schmidt W, Habel A;
 PI WPI; 2004-169243/16.
 DR New polypeptide encoded by an alternative reading frame of a pathogenic
 PT virus comprising an antigenic determinant, useful for treating or
 PT preventing an infection with the pathogenic virus.
 XX Claim 18; Page 199; 220pp; English.
 XX This invention relates to a novel polypeptide encoded by an alternative
 CC reading frame of a pathogenic virus, where the polypeptide starts with a
 CC methionine amino acid residue, which comprises an antigenic determinant
 CC and more than 7 amino acid residues. The invention may be useful for the
 CC production of compounds with a virucide activity or the development of a
 CC vaccine. The polypeptide or its fragments may be useful as a therapeutic
 CC agent. It is also useful for the manufacture of a medicament for treating
 CC or preventing an infection with the pathogenic virus. The present
 CC sequence is that of a human papillomavirus (HPV) epitope peptide of the
 CC invention.
 XX Sequence 9 AA;
 SQ

Query Match 92.9%; Score 39; DB 8; Length 9;
 Best Local Similarity 85.7%; Pred. No. 1.8e+06;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 MQPPPLP 7
 Db 1 LQPPPLP 7

RESULT 14
 ADK10337
 ID ADK10337 standard; peptide; 10 AA.
 XX
 AC ADK10337;
 XX
 DT 06-MAY-2004 (first entry)
 XX
 DE Human papillomavirus peptide #2392.
 XX
 KW pathogenic virus; alternative reading frame; antigenic determinant;
 KW virucide; vaccine; therapeutic agent; infection; HPV.
 XX
 OS Human papillomavirus.
 XX
 PN WO2004011650-A2.
 XX
 PD 05-FEB-2004.
 XX
 PF 24-JUL-2003; 2003WO-EP008112.
 XX
 PR 24-JUL-2002; 2002AT-00001124.
 XX
 PR 11-JUL-2003; 2003EP-00450171.
 XX
 PA (INTE-) INTERCELL AG.
 XX
 PI Mattner F, Schmidt W, Habel A;
 XX
 PS WPI; 2004-169243/16.
 XX
 PT New polypeptide encoded by an alternative reading frame of a pathogenic
 PT virus comprising an antigenic determinant, useful for treating or
 PT preventing an infection with the pathogenic virus.
 XX Claim 18; Page 199; 220pp; English.
 XX This invention relates to a novel polypeptide encoded by an alternative
 CC reading frame of a pathogenic virus, where the polypeptide starts with a
 CC methionine amino acid residue, which comprises an antigenic determinant
 CC and more than 7 amino acid residues. The invention may be useful for the

CC production of compounds with a virucide activity or the development of a
 CC vaccine. The polypeptide or its fragments may be useful as a therapeutic
 CC agent. It is also useful for the manufacture of a medicament for treating
 CC or preventing an infection with the pathogenic virus. The present
 CC sequence is that of a human papillomavirus (HPV) epitope peptide of the
 CC invention.
 XX Sequence 10 AA;
 SQ

Query Match 92.9%; Score 39; DB 8; Length 10;
 Best Local Similarity 85.7%; Pred. No. 45;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 MQPPPLP 7
 Db 1 LQPPPLP 7

RESULT 15
 ADK10338
 ID ADK10338 standard; peptide; 10 AA.
 XX
 AC ADK10338;
 XX
 DT 06-MAY-2004 (first entry)
 XX
 DE Human papillomavirus peptide #2393.
 XX
 KW pathogenic virus; alternative reading frame; antigenic determinant;
 KW virucide; vaccine; therapeutic agent; infection; HPV.
 XX
 OS Human papillomavirus.
 XX
 PN WO2004011650-A2.
 XX
 PD 05-FEB-2004.
 XX
 PF 24-JUL-2003; 2003WO-EP008112.
 XX
 PR 24-JUL-2002; 2002AT-00001124.
 XX
 PR 11-JUL-2003; 2003EP-00450171.
 XX
 PA (INTE-) INTERCELL AG.
 XX
 PI Mattner F, Schmidt W, Habel A;
 XX
 PS WPI; 2004-169243/16.
 XX
 PT New polypeptide encoded by an alternative reading frame of a pathogenic
 PT virus comprising an antigenic determinant, useful for treating or
 PT preventing an infection with the pathogenic virus.
 XX Claim 18; Page 199; 220pp; English.
 XX This invention relates to a novel polypeptide encoded by an alternative
 CC reading frame of a pathogenic virus, where the polypeptide starts with a
 CC methionine amino acid residue, which comprises an antigenic determinant
 CC and more than 7 amino acid residues. The invention may be useful for the
 CC production of compounds with a virucide activity or the development of a
 CC vaccine. The polypeptide or its fragments may be useful as a therapeutic
 CC agent. It is also useful for the manufacture of a medicament for treating
 CC or preventing an infection with the pathogenic virus. The present
 CC sequence is that of a human papillomavirus (HPV) epitope peptide of the
 CC invention.
 XX Sequence 10 AA;
 SQ

Query Match 92.9%; Score 39; DB 8; Length 10;
 Best Local Similarity 85.7%; Pred. No. 45;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 MQPPPLP 7
 Db 1 LQPPPLP 7

Tue Apr 12 09:04:54 2005

us-10-691-330-1.rag

Page 8

Db 1 LQPPPLP 7

Search completed: April 12, 2005, 08:18:13
Job time : 42.2 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 12, 2005, 07:49:30 ; Search time 8.20909 Seconds
(without alignments)
63.654 Million cell updates/sec

Title: US-10-691-330-1

Perfect score: 42

Sequence: 1 MQPPPLP 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5A COMB.pap:*
2: /cgn2_6/ptodata/1/iaa/5B COMB.pap:*
3: /cgn2_6/ptodata/1/iaa/6A COMB.pap:*
4: /cgn2_6/ptodata/1/iaa/6B COMB.pap:*
5: /cgn2_6/ptodata/1/iaa/PCTUS COMB.pap:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	42	100.0	7	4	US-09-641-803-1
2	42	100.0	147	4	US-09-270-767-39742
3	42	100.0	147	4	US-09-270-767-54959
4	38	90.5	452	4	US-09-270-767-43029
5	37	88.1	278	3	US-08-339-214-16
6	37	88.1	278	3	US-08-339-214-26
7	37	88.1	278	4	US-09-131-237C-6
8	37	88.1	278	4	US-09-246-129B-6
9	37	88.1	296	4	US-09-248-796A-27326
10	37	88.1	313	4	US-09-270-767-43189
11	37	88.1	479	4	US-09-252-991A-31536
12	37	88.1	573	4	US-09-248-796A-18765
13	37	88.1	642	4	US-09-252-991A-32104
14	37	88.1	649	4	US-09-538-092-1357
15	37	88.1	667	2	US-08-718-661-2
16	37	88.1	713	4	US-09-949-016-10902
17	37	88.1	721	4	US-09-533-029-78
18	37	88.1	1125	4	US-09-949-016-10194
19	36	85.7	349	1	US-08-118-270-71
20	36	85.7	349	5	PCT-US93-08528-71
21	36	85.7	705	4	US-09-902-540-11260
22	35	83.3	23	1	US-08-268-251-34
23	35	83.3	23	5	PCT-US93-01112-34
24	35	83.3	320	4	US-09-248-796A-24758
25	35	83.3	324	4	US-09-746-801A-61
26	35	83.3	383	4	US-09-248-796A-23236
27	35	83.3	447	3	US-09-388-349-9

28	35	83.3	463	4	US-10-029-180-80	Sequence 80, Appl
29	35	83.3	569	4	US-09-902-540-12184	Sequence 12184, A
30	35	83.3	623	4	US-09-949-016-6530	Sequence 6530, Ap
31	35	83.3	633	4	US-09-949-016-7806	Sequence 7806, Ap
32	35	83.3	1506	4	US-09-902-540-10944	Sequence 10944, A
33	35	83.3	2375	4	US-09-538-092-1131	Sequence 1131, Ap
34	35	83.3	2476	4	US-09-824-574-7	Sequence 7, Appli
35	34	81.0	64	4	US-09-248-796A-23598	Sequence 23598, A
36	34	81.0	65	4	US-09-248-796A-21402	Sequence 21402, A
37	34	81.0	76	4	US-09-096-731A-14	Sequence 14, Appl
38	34	81.0	111	4	US-09-902-540-14618	Sequence 14618, A
39	34	81.0	132	4	US-09-902-540-13245	Sequence 13245, A
40	34	81.0	158	4	US-09-252-991A-24956	Sequence 24956, A
41	34	81.0	194	3	US-08-822-264-4	Sequence 4, Appli
42	34	81.0	194	4	US-09-565-808-21	Sequence 21, Appl
43	34	81.0	195	2	US-08-960-022-6	Sequence 2, Appli
44	34	81.0	195	4	US-09-565-808-2	Sequence 6, Appli
45	34	81.0	223	3	US-08-822-264-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1

US-09-641-803-1

; Sequence 1, Application US/09641803

; Patent No. 6500798

; GENERAL INFORMATION:

; APPLICANT: STANTON, G. John

; APPLICANT: HUGHES, Thomas K.

; APPLICANT: BOLDOGH, Istvan

; TITLE OF INVENTION: USE OF COLOSTRININ, CONSTITUENT PEPTIDES THEREOF, AND ANALOGS THEREOF AS OXIDATIVE STRESS REGULATORS

; FILE REFERENCE: 265.00220101

; CURRENT APPLICATION NUMBER: US/09/641,803

; CURRENT FILING DATE: 2000-08-17

; PRIOR FILING DATE: 1999-08-17

; NUMBER OF SEQ ID NOS: 34

; SOFTWARE: Patentin Ver. 2.1

; SEQ ID NO 1

; LENGTH: 7

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE: Description of Artificial Sequence: synthetic

; OTHER INFORMATION: peptide

; OTHER INFORMATION: peptide

US-09-641-803-1

Query Match 100.0%; Score 42; DB 4; Length 7;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MQPPPLP 7

Db 1 MQPPPLP 7

RESULT 2

US-09-270-767-39742

; Sequence 39742, Application US/09270767

; Patent No. 6703491

; GENERAL INFORMATION:

; APPLICANT: Homburger et al.

; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster

; FILE REFERENCE: File Reference: 7326-094

; CURRENT APPLICATION NUMBER: US/09/270,767

; CURRENT FILING DATE: 1999-03-17

; NUMBER OF SEQ ID NOS: 62517

; SOFTWARE: Patentin Ver. 2.0

; SEQ ID NO 39742

; LENGTH: 147

; TYPE: PRT

```

; ORGANISM: Drosophila melanogaster
US-09-270-767-39742

Query Match      100.0%; Score 42; DB 4; Length 147;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MQPPPLP 7
DB      139 MQPPPLP 145

RESULT 3
US-09-270-767-54959
; Sequence 54959, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 54959
; LENGTH: 147
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-54959

Query Match      100.0%; Score 42; DB 4; Length 147;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MQPPPLP 7
DB      139 MQPPPLP 145

RESULT 4
US-09-270-767-43029
; Sequence 43029, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 43029
; LENGTH: 452
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-43029

Query Match      90.5%; Score 38; DB 4; Length 452;
Best Local Similarity 85.7%; Pred. No. 3e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 MQPPPLP 7
DB      287 IQPPPLP 293

RESULT 5
US-08-339-214-16
; Sequence 16, Application US/08339214
; Patent No. 6348334
; GENERAL INFORMATION:
; APPLICANT: Nagata, Shigikazu
; APPLICANT: Suda, Takashi
; APPLICANT: Takahashi, Tomoniro
; TITLE OF INVENTION: A Fas Ligand, A Fragment Thereof and DNA
; NUMBER OF SEQUENCES: 103
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch
; STREET: P.O. Box 747
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/339,214
; FILING DATE: 10-NOV-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Murphy Jr., Gerald M.
; REGISTRATION NUMBER: 28,977
; REFERENCE/DOCKET NUMBER: 1110-139P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-205-8000
; TELEFAX: 703-205-8050
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 278 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-339-214-16

Query Match      88.1%; Score 37; DB 3; Length 278;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 QPPPLP 7
DB      60 QPPPLP 65

RESULT 6
US-08-339-214-26
; Sequence 26, Application US/08339214
; Patent No. 6348334
; GENERAL INFORMATION:
; APPLICANT: Nagata, Shigikazu
; APPLICANT: Suda, Takashi
; APPLICANT: Takahashi, Tomoniro
; TITLE OF INVENTION: A Fas Ligand, A Fragment Thereof and DNA
; NUMBER OF SEQUENCES: 103
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch
; STREET: P.O. Box 747
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/339,214

```

```
; FILING DATE: 10-NOV-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Murphy Jr., Gerald M.
; REGISTRATION/DOCKET NUMBER: 28,977
; REFERENCE/DOCKET NUMBER: 1110-139P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-205-8000
; TELEFAX: 703-205-8050
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 278 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-339-214-26

Query Match      88.1%; Score 37; DB 3; Length 278;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      2 QPPPLP 7
Db      60 QPPPLP 65

RESULT 7
US-09-131-237C-6
; Sequence 6, Application US/09131237C
; Patent No. 659719
; GENERAL INFORMATION:
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Ni, Jian
; APPLICANT: Rosen, Craig A.
; TITLE OF INVENTION: Tumor Necrosis Factor Gamma
; FILE REFERENCE: PF141P3
; CURRENT APPLICATION NUMBER: US/09/131.237C
; CURRENT FILING DATE: 1998-08-07
; PRIOR APPLICATION NUMBER: 60/074,047
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: 09/005,020
; PRIOR FILING DATE: 1998-01-09
; PRIOR APPLICATION NUMBER: 08/461,246
; PRIOR FILING DATE: 1995-06-05
; PRIOR APPLICATION NUMBER: PCT/US94/12880
; PRIOR FILING DATE: 1994-11-07
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 278
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-131-237C-6

Query Match      88.1%; Score 37; DB 4; Length 278;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      2 QPPPLP 7
Db      60 QPPPLP 65

RESULT 8
US-09-246-129B-6
; Sequence 6, Application US/09246129B
; Patent No. 6824767
; GENERAL INFORMATION:
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Ni, Jian
; APPLICANT: Rosen, Craig A.
; TITLE OF INVENTION: Tumor Necrosis Factor Gamma
; FILE REFERENCE: PF141P4
```

```
; CURRENT APPLICATION NUMBER: US/09/246.129B
; CURRENT FILING DATE: 1999-02-08
; PRIOR APPLICATION NUMBER: 60/074,047
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: 09/131,237
; PRIOR FILING DATE: 1998-08-07
; PRIOR APPLICATION NUMBER: 09/005,020
; PRIOR FILING DATE: 1998-01-09
; PRIOR APPLICATION NUMBER: 08/461,246
; PRIOR FILING DATE: 1995-06-05
; PRIOR APPLICATION NUMBER: PCT/US94/12880
; PRIOR FILING DATE: 1994-11-07
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 278
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-246-129B-6

Query Match      88.1%; Score 37; DB 4; Length 278;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      2 QPPPLP 7
Db      60 QPPPLP 65

RESULT 9
US-09-248-796A-27326
; Sequence 27326, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248.796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 27326
; LENGTH: 296
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-27326

Query Match      88.1%; Score 37; DB 4; Length 296;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      2 QPPPLP 7
Db      180 QPPPLP 185

RESULT 10
US-09-270-767-43189
; Sequence 43189, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270.767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 43189
```

```
; LENGTH: 313
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-43189

Query Match      88.1%; Score 37; DB 4; Length 313;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 QPPPLP 7
DB      223 QPPPLP 228

RESULT 11
US-09-252-991A-31536
; Sequence 31536, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 31536
; LENGTH: 479
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31536

Query Match      88.1%; Score 37; DB 4; Length 479;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 QPPPLP 7
DB      193 QPPPLP 198

RESULT 12
US-09-248-796A-18765
; Sequence 18765, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 18765
; LENGTH: 573
; TYPE: PRT
; ORGANISM: Candida albicans
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (571)
; OTHER INFORMATION: Identity of amino acid sequences at the above locations are unkno
US-09-248-796A-18765

Query Match      88.1%; Score 37; DB 4; Length 573;
Best Local Similarity 100.0%; Pred. No. 5.1e+02;
```

```
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 QPPPLP 7
DB      294 QPPPLP 299

RESULT 13
US-09-252-991A-32104
; Sequence 32104, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 32104
; LENGTH: 642
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-32104

Query Match      88.1%; Score 37; DB 4; Length 642;
Best Local Similarity 100.0%; Pred. No. 5.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 QPPPLP 7
DB      41 QPPPLP 46

RESULT 14
US-09-538-092-1357
; Sequence 1357, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Giot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CurapatSeqFormatter Version 0.9
; SEQ ID NO 1357
; LENGTH: 649
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Polypeptide Accession Number Q16643
US-09-538-092-1357

Query Match      88.1%; Score 37; DB 4; Length 649;
Best Local Similarity 100.0%; Pred. No. 5.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 QPPPLP 7
DB      363 QPPPLP 368
```


RESULT 15
 US-08-718-661-2
 ; Sequence 2, Application US/08718661
 ; Patent No. 5876972
 ; GENERAL INFORMATION:
 ; APPLICANT: Nucleic acid molecules coding for mammalian
 ; TITLE OF INVENTION: tumor suppressor proteins and methods for their isolation
 ; NUMBER OF SEQUENCES: 15
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/718,661
 ; FILING DATE:
 ; CLASSIFICATION: 530
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 667 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 US-08-718-661-2

Query Match 88.1%; Score 37; DB 2; Length 667;
 Best Local Similarity 100.0%; Pred. No. 5.9e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 QPPLP 7
 Db 527 QPPLP 532

Search completed: April 12, 2005, 08:07:48
 Job time : 9.20909 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 12, 2005, 07:49:30 ; Search time 21.7 Seconds
(without alignments)
107.096 Million cell updates/sec

Title: US-10-691-330-1

Perfect score: 42

Sequence: 1 MQPPPLP 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1418010 seqs, 331997259 residues

Total number of hits satisfying chosen parameters: 1418010

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	42	100.0	7	14 US-10-281-652-1	Sequence 1, Appli
2	42	100.0	7	17 US-10-691-330-1	Sequence 1, Appli
3	39	92.9	76	15 US-10-424-599-147458	Sequence 147458,
4	39	92.9	132	9 US-09-864-761-37187	Sequence 37187, A
5	39	92.9	332	16 US-10-437-963-143948	Sequence 143948,
6	39	92.9	407	16 US-10-437-963-175374	Sequence 175374,
7	39	92.9	572	15 US-10-369-493-9267	Sequence 9267, Ap
8	39	92.9	572	15 US-10-369-493-9267	Sequence 9495, Ap
9	39	92.9	612	15 US-10-369-493-17539	Sequence 17539, A
10	38	90.5	189	15 US-10-424-599-148127	Sequence 148127,
11	38	90.5	256	16 US-10-408-765A-2896	Sequence 2896, Ap
12	38	90.5	414	15 US-10-424-599-147426	Sequence 147426,
13	38	90.5	812	16 US-10-437-963-143219	Sequence 143219,

14	37	88.1	38	10	US-09-899-495-68	Sequence 68, Appl
15	37	88.1	38	10	US-09-802-668-109	Sequence 109, App
16	37	88.1	38	14	US-10-243-475-109	Sequence 109, App
17	37	88.1	61	15	US-10-424-599-193704	Sequence 193704,
18	37	88.1	66	15	US-10-424-599-170169	Sequence 170169,
19	37	88.1	72	16	US-10-437-963-177968	Sequence 177968,
20	37	88.1	75	16	US-10-437-963-159895	Sequence 159895,
21	37	88.1	82	16	US-10-437-963-145289	Sequence 145289,
22	37	88.1	83	16	US-10-437-963-190127	Sequence 190127,
23	37	88.1	91	16	US-10-437-963-160130	Sequence 160130,
24	37	88.1	98	15	US-10-425-114-48917	Sequence 48917, A
25	37	88.1	98	16	US-10-437-963-137655	Sequence 137655,
26	37	88.1	104	16	US-10-437-963-122299	Sequence 122299,
27	37	88.1	112	16	US-10-437-963-150616	Sequence 150616,
28	37	88.1	113	16	US-10-437-963-12527	Sequence 12527,
29	37	88.1	114	15	US-10-424-599-236948	Sequence 236948,
30	37	88.1	115	16	US-10-437-963-194122	Sequence 194122,
31	37	88.1	118	16	US-10-437-963-196882	Sequence 196882,
32	37	88.1	129	15	US-10-424-599-192878	Sequence 192878,
33	37	88.1	139	16	US-10-767-701-51966	Sequence 51966, A
34	37	88.1	144	15	US-10-424-599-264209	Sequence 264209,
35	37	88.1	144	16	US-10-767-701-40579	Sequence 40579, A
36	37	88.1	146	15	US-10-424-599-270730	Sequence 270730,
37	37	88.1	147	16	US-10-437-963-111116	Sequence 111116,
38	37	88.1	159	16	US-10-437-963-131050	Sequence 131050,
39	37	88.1	162	16	US-10-437-963-203781	Sequence 203781,
40	37	88.1	163	16	US-10-437-963-155754	Sequence 155754,
41	37	88.1	164	16	US-10-437-963-179013	Sequence 179013,
42	37	88.1	167	16	US-10-437-963-130977	Sequence 130977,
43	37	88.1	171	16	US-10-437-963-203782	Sequence 203782,
44	37	88.1	178	16	US-10-767-701-57257	Sequence 57257, A
45	37	88.1	180	16	US-10-767-701-34818	Sequence 34818, A

ALIGNMENTS

RESULT 1

US-10-281-652-1
; Sequence 1, Application US/10281652
; Publication No. US20030091606A1
; GENERAL INFORMATION:
; APPLICANT: STANTON, G. John
; APPLICANT: HUGHES, Thomas K.
; APPLICANT: BOLDOGH, Istvan
; TITLE OF INVENTION: USE OF COLOSTRIN, CONSTITUENT PEPTIDES THEREOF, AND ANALOGS THEREOF AS OXIDATIVE STRESS REGULATORS
; FILE REFERENCE: 265.00220101
; CURRENT APPLICATION NUMBER: US/10/281.652
; PRIOR FILING DATE: 2002-10-28
; PRIOR FILING DATE: 2000-08-17
; PRIOR APPLICATION NUMBER: 60/149,310
; PRIOR FILING DATE: 1999-08-17
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: peptide
US-10-281-652-1

Query Match 100.0%; Score 42; DB 14; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.3e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MQPPPLP 7

Db 1 MQPPPLP 7

RESULT 2
 US-10-691-330-1
 ; Sequence 1, Application US/10691330
 ; Publication No. US20050042300A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Boldogh, Istvan
 ; APPLICANT: Stanton, G. John
 ; APPLICANT: Georgiades, Jerzy A.
 ; APPLICANT: Hughes, Thomas K., Jr.
 ; APPLICANT: Kruzel, Marian
 ; TITLE OF INVENTION: USE OF COLOSTRIN, CONSTITUENT PEPTIDES THEREOF, AND ANALOGS
 ; FILE REFERENCE: 265 00390101
 ; CURRENT APPLICATION NUMBER: US/10/691,330
 ; CURRENT FILING DATE: 2003-10-22
 ; PRIOR APPLICATION NUMBER: US 60/420,369
 ; PRIOR FILING DATE: 2002-10-22
 ; NUMBER OF SEQ ID NOS: 34
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 1
 ; LENGTH: 7
 ; TYPE: PRT
 ; ORGANISM: artificial
 ; FEATURE:
 ; OTHER INFORMATION: Synthetic Peptides
 US-10-691-330-1

Query Match 100.0%; Score 42; DB 17; Length 7;
 Best Local Similarity 100.0%; Pred. No. 1.3e+06;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MQPPPLP 7
 |||||
 DB 1 MQPPPLP 7

RESULT 3
 US-10-424-599-147458
 ; Sequence 147458, Application US/10424599
 ; Publication No. US20040031072A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa Thomas J
 ; APPLICANT: Kovalic David K
 ; APPLICANT: Zhou Yihua
 ; APPLICANT: Cao Yongwei
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
 ; FILE REFERENCE: 38-21(53223)B
 ; CURRENT APPLICATION NUMBER: US/10/424,599
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 285684
 ; SEQ ID NO 147458
 ; LENGTH: 76
 ; TYPE: PRT
 ; ORGANISM: Glycine max
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MRT3847_104173C.1.pap
 US-10-424-599-147458

Query Match 92.9%; Score 39; DB 15; Length 76;
 Best Local Similarity 85.7%; Pred. No. 2.2e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MQPPPLP 7
 :|||||
 DB 19 LQPPPLP 25

RESULT 4
 US-09-864-761-37187
 ; Sequence 37187, Application US/09864761
 ; Patent No. US20020048763A1

; GENERAL INFORMATION:
 ; APPLICANT: Penn, Sharron G.
 ; APPLICANT: Rank, David R.
 ; APPLICANT: Hanzel, David K.
 ; APPLICANT: Chen, Wensheng
 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
 ; FILE REFERENCE: Aesmica-X-1
 ; CURRENT APPLICATION NUMBER: US/09/864,761
 ; CURRENT FILING DATE: 2001-05-23
 ; PRIOR APPLICATION NUMBER: US 60/180,312
 ; PRIOR FILING DATE: 2000-02-04
 ; PRIOR APPLICATION NUMBER: US 60/207,456
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: US 09/632,366
 ; PRIOR FILING DATE: 2000-08-03
 ; PRIOR APPLICATION NUMBER: GB 24263.6
 ; PRIOR FILING DATE: 2000-10-04
 ; PRIOR APPLICATION NUMBER: US 60/236,359
 ; PRIOR FILING DATE: 2000-09-27
 ; PRIOR APPLICATION NUMBER: PCT/US01/00666
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00667
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00664
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00669
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00665
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00668
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00663
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00662
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00661
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00670
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: US 60/234,687
 ; PRIOR FILING DATE: 2000-09-21
 ; PRIOR APPLICATION NUMBER: US 09/608,408
 ; PRIOR FILING DATE: 2000-06-30
 ; PRIOR APPLICATION NUMBER: US 09/774,203
 ; PRIOR FILING DATE: 2001-01-29
 ; NUMBER OF SEQ ID NOS: 49117
 ; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
 ; SEQ ID NO 37187
 ; LENGTH: 132
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: MAP TO AL096814.11
 ; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.3
 ; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.8
 ; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.7
 ; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.7
 ; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.7
 ; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.3
 ; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 2.6
 ; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.7
 ; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3
 ; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.3
 ; OTHER INFORMATION: EST_HUMAN HIT: AW833678.1, EVALUAE 4.00e-03
 US-09-864-761-37187

Query Match 92.9%; Score 39; DB 9; Length 132;
 Best Local Similarity 85.7%; Pred. No. 3.6e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MQPPPLP 7
 :|||||

Db 116 LQPPPLP 122

RESULT 5

US-10-437-963-143948
; Sequence 143948, Application US/10437963
; Publication No. US20040123343A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping

; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53221)B

; CURRENT APPLICATION NUMBER: US/10/437,963

; CURRENT FILING DATE: 2003-05-14

; NUMBER OF SEQ ID NOS: 204966

; SEQ ID NO 143948

; LENGTH: 332

; TYPE: PRT

; ORGANISM: Oryza sativa

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT_MRT4530_44807C.1.pep
US-10-437-963-143948

Query Match 92.9%; Score 39; DB 16; Length 332;

Best Local Similarity 85.7%; Pred. No. 8.2e+02;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MQPPPLP 7

Db 85 LQPPPLP 91

RESULT 6

US-10-437-963-175374

; Sequence 175374, Application US/10437963

; Publication No. US20040123343A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping

; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53221)B

; CURRENT APPLICATION NUMBER: US/10/437,963

; CURRENT FILING DATE: 2003-05-14

; NUMBER OF SEQ ID NOS: 204966

; SEQ ID NO 175374

; LENGTH: 407

; TYPE: PRT

; ORGANISM: Oryza sativa

; FEATURE:

; NAME/KEY: unsure

; LOCATION: (1)..(407)

; OTHER INFORMATION: unsure at all Xaa locations

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT_MRT4530_73225C.1.pep
US-10-437-963-175374

Query Match 92.9%; Score 39; DB 16; Length 407;

Best Local Similarity 85.7%; Pred. No. 9.8e+02;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MQPPPLP 7

Db 105 LQPPPLP 111

RESULT 7

US-10-369-493-9267

; Sequence 9267, Application US/10369493

; Publication No. US20030233675A1

; GENERAL INFORMATION:

; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng

; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES

; FILE REFERENCE: 38-10(52052)B

; CURRENT APPLICATION NUMBER: US/10/369,493

; CURRENT FILING DATE: 2003-02-28

; PRIOR APPLICATION NUMBER: US 60/360,039

; PRIOR FILING DATE: 2002-02-21

; NUMBER OF SEQ ID NOS: 47374

; SEQ ID NO 9267

; LENGTH: 572

; TYPE: PRT

; ORGANISM: Xylella fastidiosa

; OTHER INFORMATION: US-10-369-493-9267

Query Match 92.9%; Score 39; DB 15; Length 572;

Best Local Similarity 85.7%; Pred. No. 1.3e+03;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MQPPPLP 7

Db 356 LQPPPLP 362

RESULT 8

US-10-369-493-9495

; Sequence 9495, Application US/10369493

; Publication No. US20030233675A1

; GENERAL INFORMATION:

; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng

; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES

; FILE REFERENCE: 38-10(52052)B

; CURRENT APPLICATION NUMBER: US/10/369,493

; CURRENT FILING DATE: 2003-02-28

; PRIOR APPLICATION NUMBER: US 60/360,039

; PRIOR FILING DATE: 2002-02-21

; NUMBER OF SEQ ID NOS: 47374

; SEQ ID NO 9495

; LENGTH: 572

; TYPE: PRT

; ORGANISM: Xylella fastidiosa

; OTHER INFORMATION: US-10-369-493-9495

Query Match 92.9%; Score 39; DB 15; Length 572;

Best Local Similarity 85.7%; Pred. No. 1.3e+03;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MQPPPLP 7

Db 356 LQPPPLP 362

RESULT 9

```

US-10-369-493-17539
; Sequence 17539, Application US/10369493
; Publication No. US2003033675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 17539
; LENGTH: 612
; TYPE: PRT
; ORGANISM: Xylella fastidiosa
US-10-369-493-17539

Query Match          92.9%; Score 39; DB 15; Length 612;
Best Local Similarity 85.7%; Pred. No. 1.4e+03;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MQPPPLP 7
Db 356 LQPPPLP 362

RESULT 10
US-10-424-599-148127
; Sequence 148127, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: SOY Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 148127
; LENGTH: 189
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(189)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_104780C.1.pep
US-10-424-599-148127

Query Match          90.5%; Score 38; DB 15; Length 189;
Best Local Similarity 85.7%; Pred. No. 6.7e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MQPPPLP 7
Db 105 MKPPPLP 111

RESULT 11
US-10-408-765A-2896
; Sequence 2896, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2896
; LENGTH: 256
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-2896

Query Match          90.5%; Score 38; DB 16; Length 256;
Best Local Similarity 85.7%; Pred. No. 8.8e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MQPPPLP 7
Db 191 VQPPPLP 197

RESULT 12
US-10-424-599-147426
; Sequence 147426, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: SOY Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-31(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 147426
; LENGTH: 414
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_104144C.1.pep
US-10-424-599-147426

Query Match          90.5%; Score 38; DB 15; Length 414;
Best Local Similarity 85.7%; Pred. No. 1.4e+03;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MQPPPLP 7
Db 105 MKPPPLP 111

RESULT 13
US-10-437-963-143219
; Sequence 143219, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad

```

; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21153211B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 143219
; LENGTH: 812
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_44149C.1.pep
US-10-437-963-143219

Query Match 90.5%; Score 38; DB 16; Length 812;
Best Local Similarity 85.7%; Pred. No. 2.5e+03;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MQPPPLP 7
Db 26 VQPPPLP 32

RESULT 14

US-09-899-495-68
; Sequence 68, Application US/09899495
; Publication No. US2003008060A1
; GENERAL INFORMATION:
; APPLICANT: Benjamin, Christopher W.
; APPLICANT: Roberts, Steven L.
; APPLICANT: Karnovsky, Alla M.
; APPLICANT: Ruble, Cara L.
; TITLE OF INVENTION: Human Ion Channels
; FILE REFERENCE: 00188US1
; CURRENT APPLICATION NUMBER: US/09/899,495
; CURRENT FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 60/215,815
; PRIOR FILING DATE: 2000-07-05
; PRIOR APPLICATION NUMBER: 60/216,481
; PRIOR FILING DATE: 2000-07-06
; PRIOR APPLICATION NUMBER: 60/216,479
; PRIOR FILING DATE: 2000-07-06
; PRIOR APPLICATION NUMBER: 60/216,482
; PRIOR FILING DATE: 2000-07-06
; PRIOR APPLICATION NUMBER: 60/217,096
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 125
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 68
; LENGTH: 38
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-899-495-68

Query Match 88.1%; Score 37; DB 10; Length 38;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 QPPPLP 7
Db 25 QPPPLP 30

RESULT 15

US-09-802-668-109
; Sequence 109, Application US/09802668
; Publication No. US20030190714A1
; GENERAL INFORMATION:
; APPLICANT: Wood, Linda
; APPLICANT: Vogeli, Gabriel
; APPLICANT: Karnovsky, Alla
; APPLICANT: Linske-O'Connell, Lisa I.

; APPLICANT: Wang, Jun
; APPLICANT: Liu, Derong
; TITLE OF INVENTION: Human Ion Channels
; FILE REFERENCE: 00069US1
; CURRENT APPLICATION NUMBER: US/09/802,668
; CURRENT FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: 60/188,517
; PRIOR FILING DATE: 2000-03-10
; NUMBER OF SEQ ID NOS: 117
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 109
; LENGTH: 38
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-802-668-109

Query Match 88.1%; Score 37; DB 10; Length 38;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 QPPPLP 7
Db 25 QPPPLP 30

Search completed: April 12, 2005, 07:55:20
Job time : 22.7 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 12, 2005, 07:49:31 ; Search time 5.02727 Seconds
(without alignments)
133.973 Million cell updates/sec

Title: US-10-691-330-1

Perfect score: 42

Sequence: 1 MQPPPLP 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 79: *
1: Pir1.*
2: Pir2.*
3: Pir3.*
4: Pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40	95.2	1776	2 G86280	protein T5E21.13 [
2	39	92.9	612	2 E82756	beta-galactosidase
3	38	90.5	536	2 T02441	DNA-(apurinic or a
4	38	90.5	542	2 H86239	protein P20B24.8 [
5	37	88.1	164	2 T15525	hypothetical prote
6	37	88.1	178	2 T00644	hypothetical prote
7	37	88.1	186	2 T22685	hypothetical prote
8	37	88.1	211	2 S55129	transcription fact
9	37	88.1	278	2 A49266	faa ligand - rat
10	37	88.1	393	2 T00575	probable WRKY-type
11	37	88.1	471	2 T21349	hypothetical prote
12	37	88.1	483	2 A25896	beta-adrenergic re
13	37	88.1	588	2 E36633	probable Serine/Th
14	37	88.1	620	2 S06733	hydroxyproline-ric
15	37	88.1	649	2 JN0809	drebrin E (clone g
16	37	88.1	651	2 T14763	hypothetical prote
17	37	88.1	731	2 JC7701	ARHGAP9 protein -
18	37	88.1	749	2 G86186	hypothetical prote
19	37	88.1	830	2 T18860	hypothetical prote
20	37	88.1	1124	2 UX0293	zinc finger protei
21	37	88.1	1125	2 T19193	hypothetical prote
22	37	88.1	1154	2 A56242	E-box-binding repr
23	37	88.1	1291	2 T17242	hypothetical prote
24	37	88.1	1364	2 T00250	MEGF2 protein - hu
25	37	88.1	1560	2 T00080	hypothetical prote
26	36	85.7	489	2 S23410	FUN19 protein - ye
27	36	85.7	505	2 S72273	actin-depolymerizi
28	36	85.7	542	2 JC7391	Ca2+-binding prote
29	36	85.7	610	2 T22687	hypothetical prote

30	36	85.7	736	2 I51691	dishevelled homolo
31	36	85.7	795	2 T49835	hypothetical prote
32	35	85.7	823	2 A36378	probable transcrip
33	35	83.3	267	2 T47307	hypothetical prote
34	35	83.3	300	2 T25119	hypothetical prote
35	35	83.3	312	2 A86279	protein F1K23.14 [
36	35	83.3	321	2 G86411	protein F1K23.14 [
37	35	83.3	337	2 F96507	hypothetical prote
38	35	83.3	414	2 A88485	protein F23F12.6 [
39	35	83.3	426	2 T47365	hypothetical prote
40	35	83.3	486	2 T21481	hypothetical prote
41	35	83.3	494	2 T19550	hypothetical prote
42	35	83.3	548	2 S52735	Cw17R protein - mo
43	35	83.3	587	2 S49942	hypothetical prote
44	35	83.3	597	2 S72468	probable transcrip
45	35	83.3	639	2 G02919	transcription fact

ALIGNMENTS

RESULT 1

G86280

protein T5E21.13 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004

C:Accession: G86280

R:Theologos, A.; Ecker, J.R.; Palm, C.J.; Federapfel, N.A.; Kaul, S.; White, O.; Alonso,

Chin, C.W.; Hughes, B.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;

ansen, N.F.; Chung, M.K.; Huizar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: G86280

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1776 <STO>

A:Cross-references: UNIPROT:Q9NA20; GB:AE005172; NID:97527720; PIDN:AAF63169.1; GSPDB:GN

C:Genetics:

A:Gene: T5E21.13

A:Map position: 1

Query Match 95.2%; Score 40; DB 2; Length 1776;

Best Local Similarity 85.7%; Pred. No. 2.1e+02;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MQPPPLP 7

Db 1643 MQPPMP 1649

RESULT 2

E82756

beta-galactosidase X70840 [imported] - Xylella fastidiosa (strain 9a5c)

C:Species: Xylella fastidiosa

C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004

C:Accession: E82756

R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen

Nature 406, 151-157, 2000

A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.

A:Reference number: A82515; MUID:20365717; PMID:10910347

A>Note: for a complete list of authors see reference number A59328 below

A:Accession: E82756

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-612 <STM>

A:Cross-references: UNIPROT:Q9PF38; GB:AE003923; GB:AE003849; NID:99105736; PIDN:AAF8365

A:Experimental source: strain 9a5c

R.; Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H. as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S. submitted to GenBank, June 2000

A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Praga, J.S.; Franca, S.C.; Franco, M.C.; Frohm J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigz chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E. A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.; F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A. Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir M.; Teuhako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z A;Reference number: A59328

A;Contents: annotation
A;Genetics:
A;Gene: XF0840
A;Superfamily: beta-galactosidase bga

Query Match 92.9%; Score 39; DB 2; Length 612;
Best Local Similarity 85.7%; Pred. No. 93;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MQPPPLP 7
:|||||
Db 356 LQPPPLP 362

RESULT 3

T02441
DNA-(apurinic or apyrimidinic site) lyase (ARP) [imported] - Arabidopsis thaliana
N;Alternate names: hypothetical protein T26J13.5

C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 09-Jul-2004
C;Accession: T02441; A84842

R;Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul submitted to the EMBL Data Library, June 1998

A;Description: Arabidopsis thaliana chromosome II BAC T26J13 genomic sequence.
A;Reference number: Z14673

A;Accession: T02441
A;Status: translated from GB/EMBL/DBJ

A;Molecule type: DNA
A;Residues: 1-536 <ROU>

A;Cross-references: UNIPROT:P45951; EMBL:AC004625; NID:g3241939; PID:g3241944
A;Experimental source: cultivar Columbia

R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umavam, L.; Tallon, L. euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J Nature 402, 761-768, 1999

A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197

A;Accession: A84842
A;Status: preliminary

A;Molecule type: DNA
A;Residues: 1-536 <STO>

A;Cross-references: GB:AF002093; NID:g3241944; PIDN:AAC23731.1; GSPDB:GN00139
C;Genetics:

A;Gene: AC2941460; T26J13.5
A;Map position: 2

A;Introns: 18/2; 78/3; 110/2; 137/1; 210/1; 317/3; 354/3; 400/3; 416/3; 443/3; 497/1

Query Match 90.5%; Score 38; DB 2; Length 536;

Best Local Similarity 85.7%; Pred. No. 1.2e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MQPPPLP 7
:|||||
Db 264 MRPPPLP 270

RESULT 4

H86239
proteins F20824.8 [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C;Accession: H86239

R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; ansen, N.F.; Hughes, B.; Huizar, L.

A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I. ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712

A;Accession: H86239
A;Status: preliminary

A;Molecule type: DNA
A;Residues: 1-542 <STO>

A;Cross-references: UNIPROT:Q9SGY5; GB:AB005172; NID:g6573750; PIDN:AAF17670.1; GSPDB:GN C;Genetics:

A;Gene: F20B24.8
A;Map position: 1

Query Match 90.5%; Score 38; DB 2; Length 542;

Best Local Similarity 85.7%; Pred. No. 1.2e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MQPPPLP 7
:|||||
Db 105 VQPPPLP 111

RESULT 5

T15525
hypothetical protein C16B8.3 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C;Accession: T15525

R;Bentley, D.
submitted to the EMBL Data Library, November 1995

A;Description: The sequence of C. elegans cosmid C16B8.
A;Reference number: Z18365

A;Accession: T15525
A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA
A;Residues: 1-164 <BEN>

A;Cross-references: UNIPROT:Q18043; EMBL:U41031; NID:g1098982; PID:g1098985; PIDN:AAA826 C;Genetics:

A;Gene: CESP:C16B8.3
A;Superfamily: proline-rich protein

Query Match 88.1%; Score 37; DB 2; Length 164;

Best Local Similarity 100.0%; Pred. No. 51;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 QPPPLP 7
:|||||
Db 97 QPPPLP 102

RESULT 6

T00644
hypothetical protein F316.7 - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 09-Jul-2004
C;Accession: T00644

R;Federspiel, N.A.; Palm, C.J.; Conway, A.B.; Kurtz, D.B.; Conway, A.R.; Au, M.; Araujo, ; Vysotskaia, V.S.; Yu, G.; Ecker, J.; Theologis, A.; Davis, R.W.

submitted to the EMBL Data Library, February 1998
A;Reference number: Z14197

A;Accession: T00644
A;Status: translated from GB/EMBL/DBJ

A;Molecule type: DNA
A;Residues: 1-178 <FED>

A;Cross-references: UNIPROT:O48681; EMBL:AC002396; NID:g2749918; PID:g2829866; GSPDB:GN00139
C;Genetics:
A;Gene: ATSP-F316.7
A;Map position: 1
A;Introns: 50/3; 87/1
C;Superfamily: Arabidopsis thaliana hypothetical protein F316.7

Query Match 88.1%; Score 37; DB 2; Length 178;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 QPPPLP 7
Db 162 QPPPLP 167

RESULT 7
T22685
hypothetical protein F55A11.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T22685
R;Kershaw, J.
submitted to the EMBL Data Library, May 1996
A;Reference number: Z19600
A;Accession: T22685
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-186 <WIL>
A;Cross-references: UNIPROT:Q20796; EMBL:Z72511; PIDN:CAA96655.1; GSPDB:GN00023; CESP:F55A11
A;Experimental source: clone F55A11
C;Genetics:
A;Gene: CESP:F55A11.1
A;Map position: 5
A;Introns: 33/3; 59/2; 146/2

Query Match 88.1%; Score 37; DB 2; Length 186;
Best Local Similarity 100.0%; Pred. No. 58;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 QPPPLP 7
Db 127 QPPPLP 132

RESULT 8
S55129
transcription factor RGM1 - yeast (Saccharomyces cerevisiae)
N;Alternate names: protein YN8010.12c; protein YMR182c
C;Species: Saccharomyces cerevisiae
C;Date: 08-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 09-Jul-2004
C;Accession: S55129; S17249
R;Churcher, C.M.
submitted to the EMBL Data Library, June 1995
A;Reference number: S55118
A;Accession: S55129
A;Molecule type: DNA
A;Residues: 1-211 <CHU>
A;Cross-references: UNIPROT:Q00453; EMBL:Z49808; NID:g854440; PID:g854452; MIPS:YMR182c
A;Experimental source: strain AB972
R;Estruch, F.
Nucleic Acids Res. 19, 4873-4877, 1991
A;Title: The yeast putative transcriptional repressor RGM1 is a proline-rich zinc finger
A;Reference number: S17249; NID:92020118; PMID:1923755
A;Accession: S17249
A;Molecule type: DNA
A;Residues: 1-113, 'V', 115-211 <EST>
A;Cross-references: EMBL:X59861; NID:g4317; PIDN:CAA42521.1; PID:g4318
C;Genetics:
A;Gene: SGD:RGM1
A;Cross-references: SGD:S0004794; MIPS:YMR182c
A;Map position: 13R
C;Keywords: DNA binding; transcription factor; zinc finger

Query Match 88.1%; Score 37; DB 2; Length 211;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 QPPPLP 7
Db 154 QPPPLP 159

RESULT 9
A49266
fas ligand - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
C;Accession: A49266
R;Suda, T.; Takahashi, T.; Golstein, P.; Nagata, S.
Cell 75, 1169-1178, 1993
A;Title: Molecular cloning and expression of the Fas ligand, a novel member of the tumor necrosis factor gene family
A;Reference number: A49266; MUID:94084792; PMID:7505205
A;Accession: A49266
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-278 <SUD>
A;Cross-references: UNIPROT:P36940; GB:U03470; NID:g440178; PIDN:AAC52129.1; PID:g440178
C;Keywords: glycoprotein; transmembrane protein

Query Match 88.1%; Score 37; DB 2; Length 278;
Best Local Similarity 100.0%; Pred. No. 88;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 QPPPLP 7
Db 60 QPPPLP 65

RESULT 10
T00375
probable WRKY-type DNA binding protein [imported] - Arabidopsis thaliana
N;Alternate names: probable DNA-binding protein T27E13.1
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 09-Jul-2004
C;Accession: T00375; B84706
R;Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul, M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Unayam, L.; Tallon, L.; Euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J. Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: B84420; MUID:20083487; PMID:10617197
A;Accession: B84706
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-393 <ROU>
A;Cross-references: UNIPROT:O22921; EMBL:AC004165; NID:g3150396; PID:g3150397
A;Experimental source: cultivar Columbia
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Unayam, L.; Tallon, L.; Euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J. Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: B84420; MUID:20083487; PMID:10617197
A;Accession: B84706
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-393 <SUD>
A;Cross-references: GB:AE002093; NID:g2347191; PIDN:AAC16930.1; GSPDB:GN00139
C;Genetics:
A;Gene: At2g30250; T27E13.1
A;Map position: 2
A;Introns: 79/3; 117/3; 294/2; 348/2

Query Match 88.1%; Score 37; DB 2; Length 393;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 QPPPLP 7
 Db 50 QPPPLP 55

RESULT 11
 T21349
 hypothetical protein F25H2.8 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
 C:Accession: T21349
 R:Wilkinson, J.
 submitted to the EMBL Data Library, September 1996
 A:Reference number: Z19411
 A:Accession: T21349
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-471 <WIL>
 A:Cross-references: UNIPROT:Q93571; EMBL:Z79754; PIDN:CAB02096.1; GSPDB:GN00019; CESP:F25H2
 A:Experimental source: clone F25H2
 C:Genetics:
 A:Gene: CESP:F25H2.8
 A:Map position: 1
 A:Introns: 26/3; 138/3; 302/3; 353/3; 379/2
 C:Superfamily: Caenorhabditis elegans hypothetical protein F25H2.8

Query Match 88.1%; Score 37; DB 2; Length 471;
 Best Local Similarity 71.4%; Pred. No. 1.5e+02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MQPPPLP 7
 Db 451 LQPPPLP 457

RESULT 12
 A25896
 beta-adrenergic receptor - turkey
 C:Species: Meleagris gallopavo (common turkey)
 C>Date: 16-Aug-1988 #sequence_revision 16-Aug-1988 #text_change 09-Jul-2004
 C:Accession: A25896
 R:Yarden, Y.; Rodriguez, H.; Wong, S.K.F.; Brandt, D.R.; May, D.C.; Burnier, J.; Harkins
 Proc. Natl. Acad. Sci. U.S.A. 83, 6795-6799, 1986
 A:Title: The avian beta-adrenergic receptor: primary structure and membrane topology.
 A:Reference number: A25896; MUID:86313664; PMID:3018746
 A:Accession: A25896
 A:Molecule type: mRNA
 A:Residues: 1-483 <YAR>
 A:Cross-references: UNIPROT:P07700; GB:M14379; NID:G213891; PIDN:AAA49627.1; PID:G213892
 A:Experimental source: erythrocyte
 C:Superfamily: vertebrate rhodopsin
 C:Keywords: G protein-coupled receptor; transmembrane protein

Query Match 88.1%; Score 37; DB 2; Length 483;
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 QPPPLP 7
 Db 256 QPPPLP 261

RESULT 13
 E96633
 probable Serine/Threonine protein kinase F8A5.31 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
 C:Accession: E96633
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.;
 anson, N.F.; Hughes, B.; Huizart, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Accession: E96633
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-588 <STO>
 A:Cross-references: UNIPROT:O22717; GB:AE005173; NID:G2462749; PIDN:AAB71968.1; GSPDB:GN00019
 C:Genetics:
 A:Gene: F8A5.31
 A:Map position: 1

Query Match 88.1%; Score 37; DB 2; Length 588;
 Best Local Similarity 100.0%; Pred. No. 1.9e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 QPPPLP 7
 Db 550 QPPPLP 555

RESULT 14
 S06733
 hydroxyproline-rich glycoprotein precursor - common tobacco
 C:Species: Nicotiana tabacum (common tobacco)
 C>Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 09-Jul-2004
 C:Accession: S06733
 R:Keller, B.; Lamb, C.J.
 Genes Dev. 3, 1639-1646, 1989
 A:Title: Specific expression of a novel cell wall hydroxyproline-rich glycoprotein gene:
 A:Reference number: S06733; MUID:90128263; PMID:2612909
 A:Accession: S06733
 A:Molecule type: DNA
 A:Residues: 1-620 <KEL>
 A:Cross-references: UNIPROT:P13983; EMBL:X13895; NID:G19866; PIDN:CAA32090.1; PID:G19867
 C:Superfamily: hydroxyproline-rich glycoprotein
 C:Keywords: glycoprotein

Query Match 88.1%; Score 37; DB 2; Length 620;
 Best Local Similarity 100.0%; Pred. No. 2e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 QPPPLP 7
 Db 425 QPPPLP 430

RESULT 15
 JN0809
 drebrin E (clone gbbh13) - human
 C:Species: Homo sapiens (man)
 C>Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 09-Jul-2004
 C:Accession: JN0809
 R:Toda, M.; Shirao, T.; Minoshima, S.; Shimizu, N.; Toya, S.; Uyemura, K.
 Biochem. Biophys. Res. Commun. 196, 468-472, 1993
 A:Title: Molecular cloning of cDNA encoding human drebrin E and chromosomal mapping of i
 A:Reference number: JN0809; MUID:94030036; PMID:8216329
 A:Accession: JN0809
 A:Molecule type: mRNA
 A:Residues: 1-649 <TOD>
 A:Cross-references: UNIPROT:Q16643; DDBJ:D17530; NID:G498650; PIDN:BAA04480.1; PID:G49866;
 A:Experimental source: brain
 C:Comment: This protein is an actin-binding protein.
 C:Genetics:
 A:Map position: 5
 C:Keywords: actin binding

Query Match 88.1%; Score 37; DB 2; Length 649;
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 QPPPLP 7
| | | | |
Db 363 QPPPLP 368

Search completed: April 12, 2005, 07:56:46
Job time : 7.02727 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 12, 2005, 07:49:31 ; Search time 33.2182 Seconds
(without alignments)
107.909 Million cell updates/sec

Title: US-10-691-330-1
Perfect score: 42
Sequence: 1 MQPPPLP 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	42	100.0	261	Q9VAX5	Q9VAX5 drosophila
2	40	95.2	600	Q8GPV8	Q8GPV8 pseudomonas
3	40	95.2	785	Q8RFX1	Q8RFX1 arabidopsis
4	40	95.2	1776	Q9MA20	Q9MA20 arabidopsis
5	39	92.9	218	Q6C180	Q6C180 yarrowia li
6	39	92.9	222	COL1_THUOB	Q9YK2 t corticotr
7	39	92.9	315	Q75AM7	Q75AM7 ashbya goss
8	39	92.9	612	Q87AJ2	Q87AJ2 xylella fas
9	39	92.9	612	Q9PF38	Q9PF38 xylella fas
10	39	92.9	701	CG1_HUMAN	Q13495 homo sapien
11	39	92.9	1074	EM12_MOUSE	Q8K482 mus musculu
12	39	92.9	1395	Q7SC01	Q7SC01 neurospora
13	39	92.9	1404	Q9GRR6	Q9GRR6 leishmania
14	38	90.5	199	Q7QRU8	Q7QRU8 giardia lam
15	38	90.5	256	Q6Z074	Q6Z074 oryza sativ
16	38	90.5	536	ARP_ARATH	P45951 arabidopsis
17	38	90.5	542	Q9SGY5	Q9SGY5 arabidopsis
18	38	90.5	678	Q8TES7	Q8TES7 homo sapien
19	38	90.5	812	Q8LIE4	Q8LIE4 oryza sativ
20	38	90.5	2165	Q8MX0	Q8MX0 dictyosteli
21	38	90.5	2879	Q6LFQ9	Q6LFQ9 plasmodium
22	37	88.1	44	Q8SG37	Q8SG37 cyanidiosch
23	37	88.1	68	Q6ZQ3	Q6ZQ3 felis silve
24	37	88.1	72	Q87610	Q87610 chimpanzee
25	37	88.1	73	Q87603	Q87603 chimpanzee
26	37	88.1	99	Q87482	Q87482 chimpanzee
27	37	88.1	114	Q6ZQ2	Q6ZQ2 felis silve
28	37	88.1	126	Q9NDS2	Q9NDS2 penaeus jap
29	37	88.1	126	Q9NDS3	Q9NDS3 penaeus jap
30	37	88.1	128	Q8GYP2	Q8GYP2 arabidopsis
31	37	88.1	134	Q74ZR7	Q74ZR7 ashbya goss

32 37 88.1 137 2 Q84QQ0
33 37 88.1 164 2 Q18043
34 37 88.1 167 1 SERO_GALME
35 37 88.1 178 2 Q48681
36 37 88.1 184 2 Q69K83
37 37 88.1 186 2 Q20796
38 37 88.1 190 2 Q811B5
39 37 88.1 206 2 Q6ZST9
40 37 88.1 211 1 RGM1_YEAST
41 37 88.1 223 2 Q8P6X9
42 37 88.1 223 2 Q8P193
43 37 88.1 226 2 Q7XDT9
44 37 88.1 233 2 Q6YZU3
45 37 88.1 233 2 Q6ZBH3

Q84QQ0 oryza sativ
Q18043 caenorhabdi
Q76192 galliera me
Q48681 arabidopsis
Q69K83 oryza sativ
Q20796 caenorhabdi
Q811B5 mus musculu
Q6ZST9 homo sapien
Q6ZST9 saccharomyc
Q8P6X9 xanthomonas
Q8P193 xanthomonas
Q7XDT9 oryza sativ
Q6YZU3 oryza sativ
Q6ZBH3 oryza sativ

32 37 88.1 137 2 Q84QQ0
33 37 88.1 164 2 Q18043
34 37 88.1 167 1 SERO_GALME
35 37 88.1 178 2 Q48681
36 37 88.1 184 2 Q69K83
37 37 88.1 186 2 Q20796
38 37 88.1 190 2 Q811B5
39 37 88.1 206 2 Q6ZST9
40 37 88.1 211 1 RGM1_YEAST
41 37 88.1 223 2 Q8P6X9
42 37 88.1 223 2 Q8P193
43 37 88.1 226 2 Q7XDT9
44 37 88.1 233 2 Q6YZU3
45 37 88.1 233 2 Q6ZBH3

Q84QQ0 oryza sativ
Q18043 caenorhabdi
Q76192 galliera me
Q48681 arabidopsis
Q69K83 oryza sativ
Q20796 caenorhabdi
Q811B5 mus musculu
Q6ZST9 homo sapien
Q6ZST9 saccharomyc
Q8P6X9 xanthomonas
Q8P193 xanthomonas
Q7XDT9 oryza sativ
Q6YZU3 oryza sativ
Q6ZBH3 oryza sativ

ALIGNMENTS

RESULT 1

Q9VAX5 PRELIMINARY; PRT; 261 AA.
AC Q9VAX5; 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE CG12872-PA.
GN ORFNames=CG12872;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazek R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Gabor G.B.,
RA Abail J.F., Agbayani A., An H.J., Andrews-Efankoch C., Baldwin D.,
RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.V., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattet B., McIntosh T.C., McLeod M.F., McPherson D.,
RA Merklov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D., Scheeler P., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Staden-Klamis A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodgett, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yen R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.H.O.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."

```

Science 287:2185-2195(2000).
[2]
SEQUENCE FROM N.A.
RX MEDLINE=22426065; PubMed=12537568;
RA Celniker S.E., Wheeler D.A., Krommiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Friese E., Hodgson A.,
RA George R.A., Hoskins R.A., Laverty T., Wuzny D.M., Nelson C.R.,
RA Paclib J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA Svirskas R.E., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.,
RA "Finishing a whole-genome shotgun: Release 3 of the Drosophila
RT melanogaster euchromatic genome sequence."
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
[3]
SEQUENCE FROM N.A.
RX MEDLINE=22426070; PubMed=12537573;
RA Kaminker J.S., Bergman C.M., Krommiller B., Carlson J., Svirskas R.,
RA Patel S., Friese E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celniker S.E.;
RT "The transposable elements of the Drosophila melanogaster euchromatin:
a genomics perspective."
RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).
[4]
SEQUENCE FROM N.A.
RX MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
systematic review."
RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
[5]
SEQUENCE FROM N.A.
RX FlyBase;
RG FlyBase;
RN Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
[6]
SEQUENCE FROM N.A.
RX FlyBase;
RG FlyBase;
RN Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
RL EMBL; AE003763; AAF56772.2; -
DR FlyBase; FBgn0039569; CG12872.
SQ SEQUENCE 261 AA; 28401 MW; D5753D823160DA6 CRC64;

Query Match 100.0%; Score 42; DB 2; Length 261;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MQPPPLP 7
Db 194 MQPPPLP 200

RESULT 2
Q8GPF8
ID Q8GPF8 PRELIMINARY; PRT; 600 AA.
AC Q8GPF8
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE Hypothetical protein ORF SG44.
GN Name=ORF SG44;
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
[1]
RN -
RC SEQUENCE FROM N.A.
RX MEDLINE=22313472; PubMed=12426355;

```

```

DOI=10.1128/JB.184.23.6665-6680.2002;
RA Larbig K.D., Christmann A., Johann A., Klockgether J., Hartsch T.,
RA Merkl R., Wiehlmann L., Fritz H.J., Tummeler B.;
RT "Gene islands integrated into tRNA(gly) genes confer genome diversity
on a Pseudomonas aeruginosa clone."
RL J. Bacteriol. 184:6665-6680(2002).
DR EMBL; AF440524; AAM62266.1; -
DR GO; GO:0003824; F:catalytic activity; IEA.
DR InterPro; IPR011093; DUF1528.
DR InterPro; IPR003607; Met.phos hydro.
DR InterPro; IPR011119; Relax_TraI.
DR Pfam; PF07515; DUF1528; 1.
DR Pfam; PF07514; TraI_2; 1.
DR SMART; SM00471; HDG; 1.
DR Hypothetical protein.
SQ SEQUENCE 600 AA; 66371 MW; BF43C0A0E10FF9D CRC64;

Query Match 95.2%; Score 40; DB 2; Length 600;
Best Local Similarity 85.7%; Pred. No. 5.5e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MQPPPLP 7
Db 460 MQPPMP 466

RESULT 3
Q8RXF1
ID Q8RXF1 PRELIMINARY; PRT; 785 AA.
AC Q8RXF1
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Splicing factor, putative.
GN Name=Atglg14650;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
[1]
RN -
RN SEQUENCE FROM N.A.
RA Nguyen M., Karlin-Neumann G., Southwick A., Lam B., Miranda M.,
RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
RA Ecker J., Theologis A., Davis R.W.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY081293; AAL91182.1; -
DR HSP; Q862M4; 1AAR.
DR GO; GO:0003723; rRNA binding; IEA.
DR GO; GO:006396; rRNA processing; IEA.
DR InterPro; IPR000061; Surp.
DR InterPro; IPR000626; Ubiquitin.
DR Pfam; PF01805; Surp; 2.
DR Pfam; PF00240; ubiquitin; 1.
DR SMART; SM00648; SWAP; 2.
DR SMART; SM00213; UBQ; 1.
DR PROSITE; PS0128; SURP; 2.
DR PROSITE; PS0053; UBIQUITIN 2; 1.
SQ SEQUENCE 785 AA; 87536 MW; A1958C193B63BDC4 CRC64;

Query Match 95.2%; Score 40; DB 2; Length 785;
Best Local Similarity 85.7%; Pred. No. 7.3e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MQPPPLP 7
Db 652 MQPPMP 658

RESULT 4

```


Q9MA20
 ID Q9MA20 PRELIMINARY; PRT; 1776 AA.
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE T5E21.13.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Chao Q., Brooks S., Buehler E., Johnson-Hopson C., Khan S., Kim C.,
 RA Shinn P., Tambunga G., Altafi H., Bei B., Chin C., Chiu J., Choi E.,
 RA Conn L., Conway A., Gonzales A., Hansen N., Howing B., Koo T., Lam B.,
 RA Lee J., Lenz C., Li J., Liu A., Liu K., Liu S., Mukharsky N.,
 RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,
 RA Thaveri A., Toriumi M., Vaysberg M., Yu G., Federspiel N.A.,
 RA Theologis A., Ecker J.R.;
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Chao Q., Brooks S., Buehler E., Johnson-Hopson C., Khan S., Kim C.,
 RA Shinn P., Altafi H., Bei B., Chin C., Chiu J., Choi E., Conn L.,
 RA Conway A., Gonzales A., Hansen N., Howing B., Koo T., Lam B., Lee J.,
 RA Lenz C., Li J., Liu A., Liu J., Liu S., Mukharsky N., Nguyen M.,
 RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A.,
 RA Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N., Theologis A.,
 RA Ecker J.;
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Chao Q., Brooks S., Buehler E., Johnson-Hopson C., Khan S., Kim C.,
 RA Shinn P., Altafi H., Bei B., Chin C., Chiu J., Choi E., Conn L.,
 RA Conway A., Gonzales A., Hansen N., Howing B., Koo T., Lam B., Lee J.,
 RA Lenz C., Li J., Liu A., Liu J., Liu S., Mukharsky N., Nguyen M.,
 RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A.,
 RA Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N., Theologis A.,
 RA Ecker J.;
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC010657; AAF63169.1; -.
 DR PIR; G86280; G86280.
 DR HSP; Q862M4; 1AAR.
 DR GO; GO:0003723; F:RNA binding; IEA.
 DR GO; GO:0006396; P:RNA processing; IEA.
 DR InterPro; IPR000061; Surp.
 DR InterPro; IPR000626; Ubiquitin.
 DR Pfam; PF01805; Surp; 4.
 DR Pfam; PF00240; ubiquitin; 1.
 DR SMART; SM00648; SWAP; 4.
 DR SMART; SM00213; UBQ; 1.
 DR PROSITE; PS0128; SURP; 4.
 DR PROSITE; PS50053; UBIQUITIN 2; 1.
 SQ SEQUENCE 1776 AA; 200270 MW; C1C28169934E0A27 CRC64;
 Query Match 95.2%; Score 40; DB 2; Length 1776;
 Best Local Similarity 85.7%; Pred. No. 1.7e+03;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MQPPPLP 7
 DB 1643 MQPPMP 1649
 RESULT 5
 ID Q6C180 PRELIMINARY; PRT; 218 AA.
 AC Q6C180;
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)

DE Similarity.
 GN ORFNames=YAL10F185249;
 OS Yarrowia lipolytica CLIB99.
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Dipodascaceae; Yarrowia.
 OX NCBI_TaxID=284591;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
 RA Lafontaine I., de Montigny J., Marck C., Neuveglise C., Talla E.,
 RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,
 RA Barnay S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,
 RA Boisrame A., Boyer J., Cattolico L., Confaniolieri F., de Daruvar A.,
 RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
 RA Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
 RA Kerret A., Kosul R., Lemaire M., Lesur I., Ma L., Muller H.,
 RA Nicaud J.M., Nikolski M., Oztas S., Ozier-Kalogeropoulos O.,
 RA Pellenz S., Potier S., Richard G.F., Straub M.L., Suleau A.,
 RA Swennene D., Tekala F., Wesolowski-Louvel M., Westhof E., Wirth B.,
 RA Zenou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,
 RA Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,
 RA Winkler P., Souciet J.L.;
 RL "Genome evolution in yeasts."
 RN Nature 430:35-44(2004).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA STRAIN=CLIB99;
 RC Genoscope;
 RA Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; CR382132; CAG78391.1; -.
 SQ SEQUENCE 218 AA; 24159 MW; A859A029FB99BC4B CRC64;
 Query Match 92.9%; Score 39; DB 2; Length 218;
 Best Local Similarity 85.7%; Pred. No. 2.6e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MQPPPLP 7
 DB 18 MEPPPLP 24
 RESULT 6
 ID COLI_THUOB STANDARD; PRT; 222 AA.
 AC Q9YKZ2;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Corticotropin-lipotropin precursor (pro-opiomelanocortin) (POMC)
 DE (Contains: Corticotropin (Adrenocorticotrophic hormone) (ACTH);
 DE Melanotropin alpha (Alpha-MSH); Corticotropin-like intermediary
 DE peptide (CLIP); Lipotropin beta (Beta-LPH); Lipotropin gamma (Gamma-
 DE LPH); Melanotropin beta (Beta-MSH); Beta-endorphin; Met-enkephalin).
 OS Thunnus obesus (Bigeye tuna).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Perciformes; Scombroidei;
 OC Scombridae; Thunnus.
 OX NCBI_TaxID=8241;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Pituitary;
 RA Amemiya Y., Takahashi A., Kawauchi H.;
 RL "Tuna proopiomelanocortin cDNA."
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
 CC -!- PTM: Specific enzymatic cleavages at paired basic residues yield
 CC the different active peptides.
 CC -!- SIMILARITY: Belongs to the POMC family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; AB020971; BAA35125.1; -
 DR InterPro: IPR001941; Mcortin ACTH.
 DR Pfam; PF00976; ACTH_Gdomain; 1.
 DR PRINTS; PR00383; MELANOCORTIN.
 DR PRODOM; PD003250; Mcortin ACTH; 1.
 DR Cleavage on pair of basic residues; Endorphin; Hormone; Signal.
 FT SIGNAL 1 18 Potential.
 FT PEPTIDE 93 132 Corticotropin.
 FT PEPTIDE 93 107 Melanotropin alpha.
 FT PEPTIDE 111 132 Corticotropin-like intermediary peptide.
 FT PEPTIDE 136 222 Lipotropin beta.
 FT PEPTIDE 136 188 Lipotropin gamma.
 FT PEPTIDE 172 188 Melanotropin beta.
 FT PEPTIDE 191 222 Beta-endorphin.
 FT PEPTIDE 191 195 Met-enkephalin.
 FT DOMAIN 71 75 Poly-Ser.
 FT DOMAIN 80 87 Poly-Ser.
 SQ SEQUENCE 222 AA; 24970 MW; 7868C713D6360B70 CRC64;

Query Match 92.9%; Score 39; DB 1; Length 222;
 Best Local Similarity 85.7%; Pred. No. 2.7e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MQPPPLP 7
 :|||||
 Db 62 LQPPPLP 68

RESULT 7
 ID Q75AM7 PRELIMINARY; PRT; 315 AA.
 AC Q75AM7;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE ADL104Wp.
 GN ORFNames=ADL104W;
 OS Ashbya gossypii (Yeast) (Eremothecium gossypii).
 CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Saccharomycetaceae; Eremothecium.
 CC NCBI_TaxID=33169;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 10895;
 RA Voegeli S.E., Brachat S., Dietrich F.S., Lerch A., Gaffney T.,
 RA Philippsen P.;
 RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: Belongs to the bZIP family.
 DR EMBL; AE016817; AAS51816.1; -
 DR AGD; ADL104W; -
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0003677; F:DNA binding; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro; IPR008917; Euk.transcr_DNA.
 DR InterPro; IPR004827; TF_BZIP.
 DR Pfam; PF00170; bZIP_1; 1.
 DR SMART; SM00338; BRLZ; 1.
 DR PROSITE; PS0217; BZIP; 1.
 DR PROSITE; PS00036; BZIP_BASIC; 1.
 KW DNA-binding; Nuclear protein.
 SQ SEQUENCE 315 AA; 34594 MW; 27AED42A1046A7B2 CRC64;

Query Match 92.9%; Score 39; DB 2; Length 315;
 Best Local Similarity 85.7%; Pred. No. 3.9e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MQPPPLP 7

Db 101 LQPPPLP 107
 :|||||
 RESULT 8
 ID Q87AJ2 PRELIMINARY; PRT; 612 AA.
 AC Q87AJ2;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Beta-galactosidase.
 GN Name=Bga; OrderedLocNames=PD1833;
 OS Xylella fastidiosa (strain Temecual / ATCC 700964).
 CC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
 CC Xanthomonadaceae; Xylella.
 CC NCBI_TaxID=183190;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22421331; PubMed=12533478;
 RX DOI=10.1128/JB.185.3.1018-1026.2003;
 RA Van Sluys M.A., de Oliveira M.C., Monteiro-Vitorello C.B., Moon D.H.,
 RA Miyaki C.Y., Furlan L.R., Camargo L.E.A., da Silva A.C.R., da Silva F.R.,
 RA Takita M.A., Lemos E.G.M., Machado M.A., Ferro M.I.T., da Silva F.R.,
 RA Goldman M.H.S., Goldman G.H., Lemos M.V.F., El-Dorri H., Teal S.M.,
 RA Carrer H., Carraro D.M., de Oliveira R.C., Nunes L.R., Siqueira W.J.,
 RA Coutinho L.L., Kimura E.T., Ferro E.S., Harakava R., Kuramae E.E.,
 RA Marino C.L., Gaglioti E., Abreu I.L., Alves L.M.C., do Amaral A.M.,
 RA Baia G.S., Blanco S.R., Brito M.S., Cannavan F.S., Celestino A.V.,
 RA da Cunha A.F., Fenille R.C., Ferro J.A., Fornighieri E.F., Kishi L.T.,
 RA Leoni S.G., Oliveira A.R., Rosa V.E. Jr., Sasaki F.T., Sena J.A.D.,
 RA de Souza A.A., Truffi D., Tsukumo F., Yanai G.M., Zaros L.G.,
 RA Civerolo E.L., Simpson A.J.G., Almeida N.F. Jr., Setubal J.C.,
 RA Kitajima J.P.;

RT "Comparative analyses of the complete genome sequences of Pierce's
 disease and citrus variegated chlorosis strains of Xylella
 fastidiosa";
 RT J. Bacteriol. 185:1018-1026(2003).
 RL -1- CATALYTIC ACTIVITY: Hydrolysis of terminal, non-reducing beta-D-
 galactose residues in beta-D-galactosides.
 CC -1- SIMILARITY: Belongs to family 35 of glycosyl hydrolases.
 DR EMBL; AE012560; AA029665.1; -
 DR GO; GO:0009341; C:beta-galactosidase complex; IEA.
 DR GO; GO:0004565; F:beta-galactosidase activity; IEA.
 DR GO; GO:0016798; F:hydrolase activity, acting on glycosyl bonds; IEA.
 DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
 DR InterPro; IPR008979; Gal_bind_like.
 DR InterPro; IPR001944; Glyco_hydro_35.
 DR Pfam; PF01301; Glyco_hydro_35; 1.
 DR PRINTS; PR00742; GLHYDRLASE35.
 DR PROSITE; PS01182; GLYCOSYL_HYDROL_F35; 1.
 KW Complete proteome; Glycosidase; Hydrolase.
 SQ SEQUENCE 612 AA; 67992 MW; 2B6663591EB556B CRC64;

Query Match 92.9%; Score 39; DB 2; Length 612;
 Best Local Similarity 85.7%; Pred. No. 7.8e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MQPPPLP 7
 :|||||
 Db 356 LQPPPLP 362

RESULT 9
 ID Q9PF38 PRELIMINARY; PRT; 612 AA.
 AC Q9PF38;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Beta-galactosidase.
 GN OrderedLocNames=Xf0840;
 OS Xylella fastidiosa.

Query Match 92.9%; Score 39; DB 2; Length 612;
 Best Local Similarity 85.7%; Pred. No. 7.8e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MQPPPLP 7
 :|||||
 Db 356 LQPPPLP 362

RESULT 9
 ID Q9PF38 PRELIMINARY; PRT; 612 AA.
 AC Q9PF38;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Beta-galactosidase.
 GN OrderedLocNames=Xf0840;
 OS Xylella fastidiosa.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
 OC Xanthomonadaceae; Xylella.
 OX NCBI_TaxID=2371;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=9a5c;
 RX MEDLINE=20365717; PubMed=10910347; DOI=10.1038/35018003;
 RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
 RA Alvaranga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
 RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
 RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carier H.,
 RA Collauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
 RA Coutinho L.B., Cristofani M., Dias-Neto E., Docena C., El-Dorry H.,
 RA Faciniani A.P., Ferreira J.J.S., Ferreira V.C.A., Ferro J.A.,
 RA Fraga J.S., Franca S.C., Franco M.C., Frohne M., Furlan L.R.,
 RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
 RA Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
 RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
 RA Lenos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
 RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
 RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
 RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
 RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
 RA Nhani A.Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
 RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
 RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
 RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
 RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
 RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
 RA da Silveira J.F., Silvestri M.L.Z., Silveira W.J., de Souza A.A.,
 RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Teuhako M.H.,
 RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
 RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
 RA "The genome sequence of the plant pathogen Xylella fastidiosa";
 RL Nature 406:151-159(2000).
 CC -!- CATALYTIC ACTIVITY: Hydrolysis of terminal, non-reducing beta-D-
 CC galactose residues in beta-D-galactosides.
 CC -!- SIMILARITY: Belongs to family 35 of glycosyl hydrolases.
 CC EMBL; AE003923; AAF83650.1; --
 DR PIR; B82756; E82756.
 DR GO; GO:0009341; C:beta-galactosidase complex; IEA.
 DR GO; GO:0004565; F:beta-galactosidase activity; IEA.
 DR GO; GO:0016798; F:hydrolyase activity, acting on glycosyl bonds; IEA.
 DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
 DR InterPro; IPR008979; Gal bind like.
 DR InterPro; IPR001944; Glyco_hydro_35.
 DR Pfam; PF01301; Glyco_hydro_35; 1.
 DR PRINTS; PR00742; GLYDRASE35.
 DR PROSITE; PS01182; GLYCOSYL HYDROL F35; 1.
 DR Complete proteome; Glycosylase; Hydrolase.
 SQ SEQUENCE 612 AA; 67893 MW; B894D6A3EB789BCB CRC64;
 Query Match 92.9%; Score 39; DB 2; Length 612;
 Best Local Similarity 85.7%; Pred. No. 7.8e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 YQ 1 MQPPPLP 7
 Db 356 LQPPPLP 362
 RESULT 10
 CG1_HUMAN
 ID CG1_HUMAN STANDARD; PRT; 701 AA.
 AC Q13495;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE CG1 protein (F18).
 GN Name=CG1; Synonyms=CG1;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96225444; PubMed=8640223;
 RA Laporte J., Hu L.-J., Kretz C., Mandel J.-L., Kioschis P., Coy J.,
 RA Klauack S.M., Poutska A., Dahl N.;
 RT "A gene mutated in X-linked myotubular myopathy defines a new putative
 RT tyrosine phosphatase family conserved in yeast";
 RL Nat. Genet. 13:175-182(1996).
 CC -!- TISSUE SPECIFICITY: Preferentially expressed in skeletal muscle.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC EMBL; U46023; AAC50551.1; --
 DR EMBL; HGNC:2568; CXorf6.
 DR MIM; 300120; --
 FT DOMAIN 275 286 Poly-Pro.
 FT DOMAIN 419 429 Poly-Gln.
 FT DOMAIN 522 533 Poly-Gln.
 SQ SEQUENCE 701 AA; 74475 MW; AD1C18E4F6DFD34E CRC64;
 Query Match 92.9%; Score 39; DB 1; Length 701;
 Best Local Similarity 85.7%; Pred. No. 9e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 YQ 1 MQPPPLP 7
 Db 273 LQPPPLP 279
 RESULT 11
 EM12_MOUSE
 ID EM12_MOUSE STANDARD; PRT; 1074 AA.
 AC Q8K482;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 25-JAN-2005 (Rel. 46, Last annotation update)
 DE EMILIN 2 precursor (Elastin microfibril interface-located protein 2)
 DE (Elastin microfibril interfacier 2) (Basilin).
 GN Name=Emilin2;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
 RC STRAIN=C57BL/6J; TISSUE=Cochlea;
 RX MEDLINE=2272101; PubMed=12837629; DOI=10.1016/S1044-7431(03)00075-7;
 RA Anna L.L., Goodyear R., Faris J.S., Jones I., Ng L., Richardson G.,
 RA Forrest D.;
 RT "An emilin family extracellular matrix protein identified in the
 RT cochlear basilar membrane";
 RL Mol. Cell. Neurosci. 23:460-472(2003).
 CC [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Limb;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalish D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
 RA "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.",
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RL [3]
 RP DEVELOPMENTAL STAGE.
 RX MEDLINE=22209197; PubMed=1221002; DOI=10.1006/dbio.2002.0764;
 RA Leinweber C., Steidl C., Schumacher N., Erhard S., Gessier M.,
 RT "Developmental expression and biochemical characterization of Emu
 family members.",
 RL Dev. Biol. 249:204-218 (2002).
 CC -!- FUNCTION: May be responsible for anchoring smooth muscle cells to
 CC elastic fibers, and may be involved not only in the formation of
 CC the elastic fiber, but also in the processes that regulate vessel
 CC assembly. Has cell adhesive capacity. Major component of the
 CC cochlear basilar membrane (BM) which may contribute to the
 CC developmental assembly or function of the BM.
 CC -!- SUBUNIT: Homotrimer associated through a moderately stable
 CC interaction of the C-terminal globular C1q domains, allowing the
 CC nucleation of the triple helix and then a further quaternary
 CC assembly to higher order polymers via intermolecular disulfide
 CC bonds (By similarity). Interacts with EMILIN.
 CC -!- SUBCELLULAR LOCATION: Extracellular matrix. Found mainly at the
 CC interface between amorphous elastin and microfibrils.
 CC -!- TISSUE SPECIFICITY: Highest levels are present in cochlea of P8
 CC pups, followed by modest levels in adult heart and lung, and much
 CC lower levels in forebrain, brainstem, cerebellum and hypothalamus.
 CC Very low levels detected in muscle, liver, kidney and eye.
 CC -!- DEVELOPMENTAL STAGE: Low levels detected in cochlea in neonatal
 CC pups at P1. Levels increased 2-fold by P5 and rose further to 16-
 CC fold at P13. Expression declined somewhat in adult mice. At E9.5,
 CC as during all stages of development, it is strongly expressed in
 CC the neural fold, the limb buds and the heart.
 CC -!- SIMILARITY: Contains 1 C1q domain.
 CC -!- SIMILARITY: Contains 1 collagenous domain.
 CC -!- SIMILARITY: Contains 1 EMI domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL; AF468645; AAH53532.1; -.
 DR EMBL; BC053753; AAH53753.1; -.
 DR MGD; MGI:2389136; Emilin2.
 DR InterPro; IPR001073; C1q.
 DR InterPro; IPR008983; TNF-like.
 DR Pfam; PF00386; C1q; 1.
 DR PROSITE; PS00871; C1q; 1.
 DR Cell adhesion; Coiled coil; Collagen; Extracellular matrix;
 KW Glycoprotein; Signal.
 FT SIGNAL 1 Potential.
 FT CHAIN 34 1074
 FT DOMAIN 48 121
 FT DOMAIN 181 218
 FT DOMAIN 259 345
 FT DOMAIN 374 394
 FT DOMAIN 533 554
 FT DOMAIN 582 620
 FT DOMAIN 783 851
 FT DOMAIN 852 913
 FT DOMAIN 914 924
 FT DOMAIN 922 1074
 FT CARBOHYD 58
 FT N-linked (GlcNAc...) (Potential).

FT CARBOHYD 463 463 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 514 514 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 586 586 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 615 615 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 744 744 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 995 995 N-linked (GlcNAc...) (Potential).
 SQ SEQUENCE 1074 AA; 117310 MW; 4B81728C83CE52A7 CRC64;
 Query Match 92.9%; Score 39; DB 1; Length 1074;
 Best Local Similarity 85.7%; Pred. No. 1.4e+03;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 MQPPPLP 7
 Db 842 LQPPPLP 848
 RESULT 12
 Q7SC01 PRELIMINARY; PRT; 1395 AA.
 AC Q7SC01;
 DT 01-MAR-2004 (T-EMBLrel. 26, Created)
 DT 01-MAR-2004 (T-EMBLrel. 26, Last sequence update)
 DT 25-OCT-2004 (T-EMBLrel. 28, Last annotation update)
 DE Predicted protein (Hypothetical protein B22K18.130).
 GN Name=NCU09482.1; Synonyms=B22K18.130;
 OS Neurospora crassa.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
 CX NCBI_TaxID=5141;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=OR74A;
 RA Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,
 RA Jaffe D., Fitzhugh W., Ma L.-J., Smirnov S., Purcell S., Rehman B.,
 RA Elkins T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M.,
 RA Qui D., Ianakiev P., Pedersen D., Nelson M., Washburne M.,
 RA Salitrennikoff C.P., Kinsey J.A., Braun E.L., Zelter A., Schulte U.,
 RA Kothe G.O., Jedd G., Mewes W., Staben C., Marcotte E., Greenberg D.,
 RA Roy A., Foley J., Naylor J., Thomann N., Barrett R., Gnerre S.,
 RA Kamal M., Kamvysselis M., Mauceli E., Bielke C., Rudd S., Frishman D.,
 RA Krystofova S., Rasmussen C., Metzberg R.L., Perkins D.D., Kroken S.,
 RA Cogoni C., Macino G., Catchside D., Li W., Pratt R.J., Osmani S.A.,
 RA Desouza C.C., Glas L., Orbach M.J., Berglund J., Voelker R.,
 RA Yarden O., Plamann M., Seiler S., Dunlap J., Radford A., Aramayo R.,
 RA Natvig D.O., Alex L.A., Mannhaupt G., Ebbole D.J., Freitag M.,
 RA Paulsen I., Sachs M.S., Lander E.S., Nusbaum C., Birren B.,
 RT "The Genome Sequence of the Filamentous Fungus Neurospora crassa.";
 RL Nature 0:0-0 (2003).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Schulte U., Aign V., Hoheisel J., Brandt P., Fartmann B., Holland R.,
 RA Nyakatura G., Mewes H.W., Mannhaupt G.;
 RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA German Neurospora genome project;
 RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
 CC -!- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 CC EMBL; AABX01000145; EAA33970.1; -.
 DR EMBL; BX842597; CAE75735.1; -.
 DR HSSP; Q92831; LJMA.
 DR InterPro; IPR001487; Bromodomain.
 DR PROSITE; PS00014; BROMODOMAIN_2; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 1395 AA; 149956 MW; CB0D84322CDA8835 CRC64;
 Query Match 92.9%; Score 39; DB 2; Length 1395;
 Best Local Similarity 85.7%; Pred. No. 1.9e+03;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy 1 MQPPPLP 7
Db 184 LQPPPLP 190

RESULT 13
Q9GRR6 PRELIMINARY; PRT; 1404 AA.
AC Q9GRR6;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein L5882.01 (Fragment).
GN Name=L5882.01;
OS Leishmania major.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Friedlin;
RX MEDLINE=98146435; PubMed=9477341;
RA Ivens A.C., Lewis S.M., Bagherzadeh A., Zhang L., Chan H.M.,
RA Smith D.F.;
RT "A physical map of the Leishmania major Friedlin genome."
RL Genome Res. 8:135-145(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Friedlin;
RA Masuy D., Purnelle B., Goffeau A., Ivens A.C., Quail M.,
RA Rajandream M.A., Barrell B.G.;
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL445943; CAC14319.1; -
KW Hypothetical protein.
FT NON_TER
SQ SEQUENCE 1404 AA; 152867 MW; 25F42AB6702B9D3C CRC64;

Query Match 92.9%; Score 39; DB 2; Length 1404;
Best Local Similarity 85.7%; Pred. No. 1.9e+03;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MQPPPLP 7
Db 1142 LQPPPLP 1148

RESULT 14
Q7QRU8 PRELIMINARY; PRT; 199 AA.
AC Q7QRU8;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE GLP 69.26179.26778.
OS Giardia lamblia ATCC 50803.
OC Eukaryota; Diplomonadida; Hexamitidae; Giardiinae; Giardia.
OX NCBI_TaxID=184922;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WB C6;
RA Morrison H.G., McArthur A.G., Adam R.D., Aley S.B., Gillin F.D.,
RA Olsen G.J., Sogin M.L.;
RT "Draft sequence of the Giardia lamblia genome."
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AACB0100128; EAA37757.1; -
DR InterPro; IPR008493; DUF775.
DR InterPro; IPR011024; G_crystallin_SF.
DR Pfam; PF05603; DUF775; 1.
SQ SEQUENCE 199 AA; 22317 MW; 1C303F8C087DAA4B CRC64;

Query Match 90.5%; Score 38; DB 2; Length 199;
Best Local Similarity 85.7%; Pred. No. 4.4e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MQPPPLP 7
Db 166 VQPPPLP 172

Query Match 90.5%; Score 38; DB 2; Length 256;
Best Local Similarity 85.7%; Pred. No. 4.4e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MQPPPLP 7
Db 42 IQPPPLP 48

RESULT 15
Q6Z074 PRELIMINARY; PRT; 256 AA.
AC Q6Z074;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein P0412D08.6 (Hypothetical protein
DE P0035F08.28).
GN Name=P0412D08.6; Synonyms=P0035F08.28;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=;
RA Sasaki T., Matsumoto T., Katayose Y.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Sasaki T., Matsumoto T., Yamamoto K.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP005499; BAD03682.1; -
DR EMBL; AP004746; BAD03407.1; -
KW Hypothetical protein.
SQ SEQUENCE 256 AA; 27545 MW; C9FB63C981F7B0CF CRC64;
```

Search completed: April 12, 2005, 08:05:35
Job time : 36.2182 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 12, 2005, 07:49:31 ; Search time 95.2 Seconds
(without alignments)
69.064 Million cell updates/sec

Title: US-10-691-330-2

Perfect score: 89

Sequence: 1 LQTPQLQVMMEPQGD 17

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq 16Dec04:*

- 1: geneseqp1980s:*
- 2: geneseqp1990s:*
- 3: geneseqp2000s:*
- 4: geneseqp2001s:*
- 5: geneseqp2002s:*
- 6: geneseqp2003as:*
- 7: geneseqp2003bs:*
- 8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	89	100.0	17	4	AAB72501 Colostrin
2	89	100.0	17	4	AAB59310 Ewe colos
3	89	100.0	17	4	AAB72247 Colostrin
4	89	100.0	17	4	AAB72533 Colostrin
5	89	100.0	17	5	AAO14578 Neural ce
6	89	100.0	17	5	AAM51037 Colostrin
7	89	100.0	17	5	AAE20229 Colostrin
8	89	100.0	17	8	ADN60296 Constitue
9	89	100.0	17	8	ADS74386 Ovine col
10	89	100.0	18	4	AAB59341 Ewe colos
11	56	62.9	11	4	AAE07185 Colostrin
12	56	62.9	12	4	AAE07195 Modified
13	51	57.3	82	7	ABO64584 Klebsiell
14	48	53.9	692	5	ABP27487 Streptoco
15	48	53.9	692	5	ABP30759 Streptoco
16	48	53.9	692	6	ABU46454 Protein e
17	48	53.9	693	6	ABU44329 Protein e
18	48	53.9	709	5	ABB55615 Lactococc
19	46	51.7	471	8	ADR96210 Novel S.
20	46	51.7	693	2	AAW80722 S. pneumo
21	46	51.7	693	4	AAU37640 Streptoco
22	46	51.7	693	4	AAW01093 CFE 96 pr
23	46	51.7	693	6	ABU00620 S. pneumo
24	46	51.7	693	6	ABU45840 Protein e
25	46	51.7	693	7	ADG73361 Streptoco

26 46 51.7 693 7 ADG73484 Streptoco
27 46 51.7 693 8 ADK47562 Streptoco
28 45 50.6 83 4 AAM15539 Peptide #
29 45 50.6 83 4 ABB34546 Peptide #
30 45 50.6 83 4 AAM28029 Peptide #
31 45 50.6 83 4 ABB29372 Peptide #
32 45 50.6 83 4 ABB19955 Protein #
33 45 50.6 83 4 AAM67729 Human bon
34 45 50.6 83 4 AAM55332 Human bra
35 45 50.6 83 4 ABB49367 Human liv
36 45 50.6 83 4 AAM03288 Peptide #
37 45 50.6 83 5 ABB37289 Human pep
38 45 50.6 249 8 ADI26093 Human pro
39 45 50.6 249 8 ADI26091 Human pro
40 45 50.6 667 2 AAW48760 BOP1 prot
41 45 50.6 704 8 ADI26095 Human pro
42 44 49.4 50 6 AAO29916 Human bet
43 44 49.4 133 2 AAY07279 Bovine sy
44 44 49.4 134 2 AAY07278 Human bet
45 44 49.4 134 8 ADP86572 Human bet

ALIGNMENTS

RESULT 1

AAB72501
ID AAB72501 standard; peptide; 17 AA.

XX AC AAB72501;

DT 09-MAY-2001 (first entry)

XX DE Colostrin peptide #2.

XX KW Dermatological; oxidative stress regulator; colostrin.

XX OS Unidentified.

XX PN WO200112650-A2.

XX PD 22-FEB-2001.

XX PF 17-AUG-2000; 2000WO-US022665.

XX PR 17-AUG-1999; 99US-0149310P.

XX PA (TEXA) UNIV TEXAS SYSTEM.

XX PI Stanton GJ, Hughes TK, Boldogh I;
XX DR WPI; 2001-218342/22.

XX PT Modulating oxidative stress level in a cell, involves contacting the cell
XX PT with an oxidative stress regulator selected from colostrin, its
XX PT constituent peptide, analog or their combinations.

XX PS Claim 6; Page 25; 48pp; English.

XX CC The present invention relates to a method for modulating the oxidative
XX CC stress level in a cell or a patient, comprising contacting the cell with,
XX CC or administering to the patient, an oxidative stress regulator selected
XX CC from colostrin, or its constituent peptide (e.g. the present peptide),
XX CC to change the level of an oxidising species in the cell. The method can
XX CC be used to treat oxidative damage to skin, by decreasing or preventing an
XX CC increase in the level of damage to a biomolecule of the patient

XX SQ Sequence 17 AA;

Query Match 100.0%; Score 89; DB 4; Length 17;

Best Local Similarity 100.0%; Pred. No. 3.2e-07;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

KW	dementia; neurodegenerative disease; Alzheimer's disease; psychosis;
KW	neurosis; infection.
XX	
OS	Synthetic.
XX	
PN	WO20011937-A2.
XX	
PD	22-FEB-2001.
XX	
PF	17-AUG-2000; 2000WO-US022818.
XX	
PR	17-AUG-1999; 99US-0149311P.
XX	
PA	(TEXA) UNIV TEXAS SYSTEM.
PA	(REGE-) REGEN THERAPEUTICS PLC.
XX	
PI	Stanton GJ, Hughes TK, Boldogh I, Georgiades J;
XX	
DR	WFI; 2001-202804/20.
XX	
PT	Inducing a cytokine and modulating an immune response, useful for
PT	treating central nervous system diseases and bacterial and viral
PT	infections, comprises administering colostrinin as an immunological
PT	regulator.
XX	
PS	Claim 1; Page 34; 50pp; English.
XX	
CC	Sequences AAB72246 - AAB72275 represent peptides derived from clostrinin,
CC	a proline rich polypeptide aggregate contained in colostrum. The peptides
CC	have immune response modulatory activity, and are capable of inducing
CC	cytokines. Colostrinin and its derived peptides are useful for inducing
CC	cytokine production, for modulating an immunological response and for
CC	inducing blood cell proliferation. The peptides are useful in the
CC	treatment of disorders of the central nervous system, neurological
CC	disorders, mental disorders, dementia, neurodegenerative diseases,
CC	Alzheimer's disease, motor neurone disease, psychosis, neurosis, chronic
CC	disorders of the immune system, bacterial and viral infections and
CC	acquired immunological deficiencies
XX	
SQ	Sequence 17 AA;
	Query Match 100.0%; Score 89; DB 4; Length 17;
	Best Local Similarity 100.0%; Pred. No. 3.2e-07;
	Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 LQTPOPLLOVMMEPGD 17
Db	1 LQTPOPLLOVMMEPGD 17
RESULT 4	
AAB72533	
ID	AAB72533 standard; peptide; 17 AA.
AC	AAB72533;
XX	
DT	09-MAY-2001 (first entry)
XX	
DE	Colostrinin peptide #2.
XX	
KW	Neuroprotective; neural cell differentiation regulator; colostrinin;
KW	colostrum.
XX	
OS	Unidentified.
XX	
PN	WO200112651-A2.
XX	
PD	22-FEB-2001.
XX	
PF	17-AUG-2000; 2000WO-US022774.
XX	
PFR	17-AUG-1999; 99US-0149633P.
XX	

PA (TEXA) UNIV TEXAS SYSTEM.
 XX Boldogh I;
 XX WPI; 2001-226545/23.
 XX
 XX Use of colostrinin, its constituent peptide or analog as a neural cell
 PT regulator, for promoting neural cell differentiation and treating damaged
 PT neural cells in a patient.
 XX
 XX Claim 6; Page 21; 35pp; English.
 XX
 XX The present invention relates to a method for promoting neural cell
 CC differentiation and treating damaged neural cells, using colostrinin and
 CC colostrinin constituent peptides (e.g. the present peptide) as a neural
 CC cell regulator. Colostrinin is a polypeptide complex found in colostrum
 XX
 XX Sequence 17 AA;
 SQ

Query Match 100.0%; Score 89; DB 4; Length 17;
 Best Local Similarity 100.0%; Pred. No. 3.2e-07;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LQTPQLLVNMEPQGD 17
 |||||
 Db 1 LQTPQLLVNMEPQGD 17
 |||||

RESULT 5
 AA014578
 ID AA014578 standard; peptide; 17 AA.
 AC AA014578;
 XX
 XX 27-MAY-2002 (first entry)
 DT
 XX
 XX Neural cell regulatory colostrinin peptide 2.
 XX
 XX Neural cell differentiation; neural cell regulator; colostrinin peptide;
 KW neural cell formation; proline-rich polypeptide aggregate; colostrum;
 KW neural cell treatment.
 XX
 XX Unidentified.
 OS
 XX
 XX Key Location/Qualifiers
 PH Modified-site 17
 FT /note= "Optional C-terminal amide"
 FT
 XX WO200213851-A1.
 XX
 XX 21-FEB-2002.
 PD
 XX 17-AUG-2000; 2000WO-US022777.
 XX
 XX 17-AUG-2000; 2000WO-US022777.
 XX
 XX (TEXA) UNIV TEXAS SYSTEM.
 XX Boldogh I, Stanton JG, Hughes TK;
 XX WPI; 2002-269152/31.
 XX
 XX Promoting cell differentiation in a patient involves use of blood cell
 PT regulator selected from colostrinin, its constituent peptide and/or
 PT analog.
 XX
 XX Claim 7; Page 21; 37pp; English.
 XX
 XX The invention comprises a method for promoting cell differentiation (e.g.
 CC neural cell differentiation). The method involves contacting cells with a
 CC neural cell regulator (i.e. a colostrinin peptide) in order to change the
 CC cells in morphology to form neural cells. Colostrinin is a proline-rich
 CC polypeptide aggregate that is present in colostrum. The method of the

CC invention is useful for promoting the differentiation of cells and for
 CC treating damaged neural cells in a patient. The present amino acid
 CC sequence represents a specifically claimed colostrinin peptide used in
 CC the method of the invention
 XX
 XX Sequence 17 AA;
 SQ

Query Match 100.0%; Score 89; DB 5; Length 17;
 Best Local Similarity 100.0%; Pred. No. 3.2e-07;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LQTPQLLVNMEPQGD 17
 |||||
 Db 1 LQTPQLLVNMEPQGD 17
 |||||

RESULT 6
 AA051037
 ID AA051037 standard; peptide; 17 AA.
 AC AA051037;
 XX
 XX 30-MAY-2002 (first entry)
 DT
 XX
 XX Colostrinin constituent peptide.
 DE
 XX
 KW Colostrinin; colostrum; immunomodulator; cardiovascular;
 KW blood cell regulator; cytokine inducer; human.
 XX
 XX Homo sapiens.
 OS
 XX
 XX Key Location/Qualifiers
 PH Modified-site 17
 FT /note= "optional C-terminal amidation"
 FT
 XX WO200213849-A1.
 XX
 XX 21-FEB-2002.
 PD
 XX 17-AUG-2000; 2000WO-US022775.
 XX
 XX 17-AUG-2000; 2000WO-US022775.
 XX
 XX (TEXA) UNIV TEXAS SYSTEM.
 XX (REGG-) REGEN THERAPEUTICS PLC.
 XX
 XX Stanton GJ, Hughes TK, Boldogh I, Georgiades J;
 XX WPI; 2002-269150/31.
 XX
 XX Modulation of blood cell proliferation in a patient involves use of blood
 PT cell regulator selected from colostrinin, its constituent peptide and/or
 PT analog.
 XX
 XX Claim 1; Page 34; 54pp; English.
 XX
 XX The present sequence is that of a colostrinin constituent peptide that is
 CC preferred for use as an immunological regulator and as a blood cell
 CC regulator in claimed methods of the invention. Methods are claimed for:
 CC inducing a cytokine in a cell by contact with an immunological regulator,
 CC where the cell is present in a cell culture, a tissue, an organ or an
 CC organism, and the cell is mammalian, including human; modulating an
 CC immune response in a cell by contact with the immunological regulator
 CC under conditions effective to induce a cytokine; modulating an immune
 CC response in a patient by administering an immunological regulator under
 CC conditions effective to induce a cytokine, where the immunological
 CC regulator is administered topically or as part of a dietary supplement,
 CC and where the immune response is specific or non specific, an interferon
 CC response or an antibody response; modulating blood cell proliferation by
 CC contacting blood cells with a blood cell regulator, where the blood cells
 CC are present in a cell culture or an organism, are mammalian or human, and
 CC where the blood cells are increased in number or differentiated; and a
 CC method for modulating blood cell proliferation in a patent. A claimed

CC cytokine-inducing composition comprises a pharmaceutical carrier and an
 CC active agent such as the present peptide. Cytokines induced by this
 CC peptide in human leucocyte cultures include interferon-gamma, tumour
 CC necrosis factor-alpha, interleukin-4, interleukin-6, interleukin-10 and
 CC interleukin-12. It was one of the best overall inducers in almost all
 CC cytokine and blood cell proliferation experiments conducted
 XX
 SQ Sequence 17 AA;

Query Match 100.0%; Score 89; DB 5; Length 17;
 Best Local Similarity 100.0%; Pred. No. 3.2e-07;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LQTPQPLQVMMEPQGD 17
 |||||
 Db 1 LQTPQPLQVMMEPQGD 17

RESULT 7
 AAEE20229
 ID AAEE20229 standard; peptide; 17 AA.

XX
 AC AAEE20229;

DT 18-JUN-2002 (first entry)

DE Colostrinin constituent peptide #2.

XX Blood cell regulator; colostrinin; constituent peptide; oxidative stress;
 KW therapy; oxidative damage; skin; aging; wound healing; cell replacement;
 KW tissue; organ; cosmetic procedure; repair; regeneration; preservation;
 KW transplantation; implantation; dermatological; vulnary.

XX Unidentified.

XX Key Location/Qualifiers
 FH Modified-site 17
 FT /note= "Optionally C-terminal amide"
 FT

XX WO200213850-A1.

XX 21-FEB-2002.

XX 17-AUG-2000; 2000WO-US022776.

XX 17-AUG-2000; 2000WO-US022776.

XX (TEXA) UNIV TEXAS SYSTEM.

XX Stanton GJ, Hughes TK, Boldogh I;

XX WPI; 2002-269151/31.

XX Composition useful for the modulation of blood cell proliferation in a
 PT patient comprises a blood cell regulator selected from colostrinin, its
 PT constituent peptide and/or analog.

XX Claim 6; Page 25; 51pp; English.

XX The invention relates to a composition which comprises a blood cell
 CC regulator selected from colostrinin, its constituent peptide and/or
 CC analogue. The invention is used for modulating the oxidative stress level
 CC in a cell e.g. mammalian or human cell present in a cell culture, tissue,
 CC organ, or organism; or for treating oxidative damage to the skin of a
 CC patient e.g. animal or human; to modulate oxidative stress during/ after
 CC a premature birth or normal birth, preventing/delaying aging in a
 CC patient, enhancing wound healing, and the reduction of side effects of
 CC cosmetic procedures. The method changes the level of an oxidising species
 CC in the cell, such as decreases or prevents increase in the level of
 CC damage to a biomolecule of the patient selected from DNA, protein and/or
 CC lipid, compared to the same conditions when the oxidative stress
 CC regulator is not present. The modulation of oxidative stress results in
 CC enhanced repair, regeneration, and replacement of cells, tissues and

CC organs (e.g. kidney, liver, pancreas, skin, and the other internal and
 CC external organs), as well as enhanced preservation of such organs for
 CC transplantation, implantation, or scientific research. The present
 CC sequence is a colostrinin constituent peptide
 XX
 SQ Sequence 17 AA;

Query Match 100.0%; Score 89; DB 5; Length 17;
 Best Local Similarity 100.0%; Pred. No. 3.2e-07;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LQTPQPLQVMMEPQGD 17
 |||||
 Db 1 LQTPQPLQVMMEPQGD 17

RESULT 8

ADN60296
 ID ADN60296 standard; peptide; 17 AA.

XX
 AC ADN60296;

XX 29-JUL-2004 (first entry)

XX Constituent peptide of colostrinin SEQ ID NO:2.

XX modulator; colostrinin; intracellular signaling molecule modulator;
 KW 4HNE-mediated lipid peroxidation down regulation; apoptosis inhibition;
 KW DNA damage; beta-amyloid; retinoic acid; cytostatic; 4HNE inhibitor;
 KW 4HNE-protein adduct formation reduction;
 KW 4HNE-mediated glutathione depletion inhibition; p53 inhibition;
 KW c-Jun NH2-terminal kinase inhibition.

XX Synthetic.

XX WO2004037851-A2.

XX 06-MAY-2004.

XX 22-OCT-2003; 2003WO-US033423.

XX 22-OCT-2002; 2002US-0420369P.

XX (TEXA) UNIV TEXAS SYSTEM.

XX (BOLD/) BOLDOGH I.

XX (STAN/) STANTON J G.

XX (GEOR/) GEORGIADIS J A.

XX (HUGH/) HUGHES T K.

XX (KRUIZ/) KRUIZEL M.

XX Boldogh I, Stanton JG, Georgiades JA, Hughes TK, Kruzel M;

XX WPI; 2004-365494/34.

XX Use of colostrinin for e.g. modulating an intracellular signaling
 PT molecule in a cell, down regulating 4HNE-mediated lipid peroxidation in a
 PT cell, inhibiting apoptosis in a cell, or protecting against DNA damage in
 PT a cell.

XX Claim 6; SEQ ID NO 2; 46pp; English.

XX The present invention describes the use of a modulator selected from
 CC colostrinin, its constituent peptide, its active analogue, and a
 CC combination of these, for modulating an intracellular signaling molecule
 CC in a cell, down regulating 4HNE-mediated lipid peroxidation in a cell,
 CC inhibiting apoptosis in a cell, protecting against DNA damage in a cell,
 CC and reducing the toxic effect of beta-amyloid or retinoic acid on a cell.
 CC The modulator has cytostatic activity, and can be used as a 4HNE
 CC inhibitor. The modulator is useful in the manufacture of a medicament for
 CC reducing 4HNE-protein adduct formation, inhibiting 4HNE-mediated
 CC glutathione depletion, inhibiting 4HNE-induced activation of p53 protein,
 CC and/or inhibiting 4HNE-induced activation of c-Jun NH2-terminal kinases.
 CC Colostrinin, or its constituent peptide or active analogue is useful for

CC inhibiting apoptosis in a cell, protecting against DNA damage in a cell,
CC or reducing the toxic effect of beta-amyloid or retinoic acid on a cell.
CC The present sequence represents a synthetic constituent peptide of
CC colostrinin, which can be used as a modulator in the present invention.
XX
SQ Sequence 17 AA;
Query Match 100.0%; Score 89; DB 8; Length 17;
Best Local Similarity 100.0%; Pred. NO. 3.2e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 LQTPQPLLQVMEPQGD 17
Db 1 LQTPQPLLQVMEPQGD 17
RESULT 9
ADS74386
ID ADS74386 standard; peptide; 17 AA.
XX
AC ADS74386;
XX
DT 16-DEC-2004 (first entry)
XX
DE Ovine colostrinin peptide.
XX
KW Colostrum; colostrinin; sheep; peptide purification.
XX
OS Ovis aries.
XX
FN WO2004081038-A1.
XX
PD 23-SEP-2004.
XX
PF 10-MAR-2004; 2004WO-GB001014.
XX
PR 11-MAR-2003; 2003GB-00005552.
XX
PR 08-MAR-2004; 2004GB-00005190.
XX
PA (REGE-) REGEN THERAPEUTICS PLC.
XX
PI Georgiades JA, Polanowski A, Wilusz T, Kruzel ML;
XX WPI; 2004-677519/66.
XX
PT Recovering peptides such as colostrinin from mammalian colostrum, by
PT mixing colostrum with alcohol to form alcohol phase containing peptides
PT and precipitate, separating alcohol phase from precipitate, and
PT recovering alcohol phase.
XX
PS Example; SEQ ID NO 1; 41pp; English.
XX
CC The present sequence is that of a peptide that can be recovered from
CC ovine colostrinin using the method of the invention. The invention
CC provides a method for the recovery of peptides (especially colostrinin)
CC from colostrum in substantially pure, biologically active form and in
CC high yield. The method involves mixing the colostrum with an alcohol to
CC form an alcohol phase containing the colostrinin and a precipitate
CC containing higher molecular weight caseins and other proteins. Best
CC results are obtained using methanol or ethanol of at least 80%, and
CC preferably up to 100% purity. The alcohol phase is then separated from
CC the precipitate, and the colostrinin is separated from the alcohol,
CC preferably by evaporation, to form a colostrinin-rich phase, which is
CC recovered. A precipitation agent, such as ammonium sulfate, may be added
CC either to the alcohol phase or, preferably, to the colostrinin-rich phase
CC to induce precipitation of the colostrinin peptides. The method is
CC generally applicable to the separation of peptides from fluids containing
CC higher molecular weight proteins, lipids, carbohydrates and/or nucleic
CC acids. In an example from the invention, the antigenic profile of
CC peptides recovered from sheep colostrum using the alcohol precipitation
CC methods was determined by ELISA using antibodies prepared against 9
CC synthetic peptides, including a peptide having the present sequence
CC (denoted antigen class A-1).

XX
SQ Sequence 17 AA;
Query Match 100.0%; Score 89; DB 8; Length 17;
Best Local Similarity 100.0%; Pred. NO. 3.2e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 LQTPQPLLQVMEPQGD 17
Db 1 LQTPQPLLQVMEPQGD 17
RESULT 10
AAB59341
ID AAB59341 standard; peptide; 18 AA.
XX
AC AAB59341;
XX
DT 21-MAR-2001 (first entry)
XX
DE Ewe colostrinin peptide fragment derived sequence #1.
XX
KW Sheep; colostrinin; proline rich polypeptide; colostrum; immune disorder;
KW central nervous system disorder; dietary supplement; beta-amyloid plaque.
XX
OS Ovis sp.
XX
FN WO200075173-A2.
XX
PD 14-DEC-2000.
XX
PF 02-JUN-2000; 2000WO-GB002128.
XX
PR 02-JUN-1999; 99GB-00012852.
XX
PA (REGE-) REGEN THERAPEUTICS PLC.
XX
PI Georgiades JA;
XX
DR WPI; 2001-071058/08.
XX
PT Peptides having an N-terminal amino acid sequence isolated from
PT colostrinin for treating e.g. disorders of the central nervous system and
PT immune system, viral and bacterial infections, and diseases characterized
PT by amyloid plaques.
XX
PS Claim 8; Page 27; 63pp; English.
XX
CC The present invention provides the sequences of a number of peptides
CC found in ewe's colostrinin. Colostrinin is the proline-rich polypeptide
CC fragment of colostrum. These peptides can be used in the treatment of
CC central nervous system disorders such as senile dementia, Parkinson's
CC disease, Alzheimer's disease, psychosis and neurosis, immune system
CC disorders such as bacterial and viral infections, to improve the
CC development of a child's immune system, as a dietary supplement, and to
CC promote the dissolution of beta-amyloid plaques
XX
SQ Sequence 18 AA;
Query Match 100.0%; Score 89; DB 4; Length 18;
Best Local Similarity 100.0%; Pred. NO. 3.4e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 LQTPQPLLQVMEPQGD 17
Db 2 LQTPQPLLQVMEPQGD 18
RESULT 11
AAE07185
ID AAE07185 standard; peptide; 11 AA.
XX
AC AAE07185;

XX 06-NOV-2001 (first entry)
 DT Colostrinin peptide 1.
 DE Colostrinin; nootropic; neuroprotective; immunomodulatory; antibacterial;
 XX Parkinson's disease; Alzheimer's disease; mental disorder; food additive;
 KW central nervous system disorder; neurodegenerative disorder; weight loss;
 KW beta-amyloid plaque; psychosis; neurosis; cachexia; motor neuron disease;
 KW acquired immunological deficiency; neurological disorder; dementia;
 KW antiviral.
 XX Unidentified.
 OS WO200155199-A1.
 XX 02-AUG-2001.
 XX 26-JAN-2001; 2001WO-GB000329.
 XX 26-JAN-2000; 2000GB-00001825.
 XX (REG-) REGEN THERAPEUTICS PLC.
 XX Georgiades JA;
 XX WPI; 2001-488775/53.
 XX Peptide useful as an interalia in the treatment of e.g. disorders of the
 PT immune system and the central nervous system comprises ten amino-terminal
 XX amino acid sequence derived from peptides present in colostrinin.
 PS Claim 1; Page 15; 40pp; English.
 XX The invention relates to colostrinin peptide fragments which are useful,
 CC inter alia, in the treatment of chronic disorders of the immune system
 CC and the central nervous system. Colostrinin peptides are used as a
 CC medicament in the treatment of neurological disorders e.g., dementia,
 CC neurodegenerative disorders e.g., Alzheimer's disease, motor neuron
 CC disease e.g., Parkinson's disease, mental disorders e.g., psychosis and
 CC neurosis, in acquired immunological deficiencies, chronic bacterial and
 CC viral infections and diseases characterised by the presence of beta-
 CC amyloid plaques and as a dietary supplement for babies, small children,
 CC adults and senile persons, who have been subjected to chemotherapy or
 CC have suffered from cachexia or weight loss due to the chronic disease.
 CC Colostrinin peptides are also used as food additives and as an auxiliary
 CC withdrawal treatment for drug addicts, after a period of detoxification
 CC and in persons dependent on stimulants. Colostrinin peptides are used to
 CC prepare antibodies and to treat emotional disturbances, e.g. emotional
 CC disturbances of psychiatric patients in a state of depression. These
 CC colostrinin peptides improves the development of immune system in a new
 CC born child and to correct the immunological deficiencies in a child. The
 CC present sequence is colostrinin peptide 1 related to the invention
 XX Sequence 11 AA;
 SQ Query Match 62.9%; Score 56; DB 4; Length 11;
 Best Local Similarity 100.0%; Pred. No. 0.039;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 5 QPLLQVMEPEQ 15
 Db 1 QPLLQVMEPEQ 11
 RESULT 12
 AAEE07195
 ID AAEE07195 standard; peptide; 12 AA.
 XX AAEE07195;
 AC AAEE07195;
 XX 06-NOV-2001 (first entry)
 DT Colostrinin peptide 1.
 DE Colostrinin; nootropic; neuroprotective; immunomodulatory; antibacterial;
 XX Parkinson's disease; Alzheimer's disease; mental disorder; food additive;
 KW central nervous system disorder; neurodegenerative disorder; weight loss;
 KW beta-amyloid plaque; psychosis; neurosis; cachexia; motor neuron disease;
 KW acquired immunological deficiency; neurological disorder; dementia;
 KW antiviral.
 XX Unidentified.
 OS WO200155199-A1.
 XX 02-AUG-2001.
 XX 26-JAN-2001; 2001WO-GB000329.
 XX 26-JAN-2000; 2000GB-00001825.
 XX (REG-) REGEN THERAPEUTICS PLC.
 XX Georgiades JA;
 XX WPI; 2001-488775/53.
 XX Peptide useful as an interalia in the treatment of e.g. disorders of the
 PT immune system and the central nervous system comprises ten amino-terminal
 XX amino acid sequence derived from peptides present in colostrinin.
 PS Claim 1; Page 15; 40pp; English.
 XX The invention relates to colostrinin peptide fragments which are useful,
 CC inter alia, in the treatment of chronic disorders of the immune system
 CC and the central nervous system. Colostrinin peptides are used as a
 CC medicament in the treatment of neurological disorders e.g., dementia,
 CC neurodegenerative disorders e.g., Alzheimer's disease, motor neuron
 CC disease e.g., Parkinson's disease, mental disorders e.g., psychosis and
 CC neurosis, in acquired immunological deficiencies, chronic bacterial and
 CC viral infections and diseases characterised by the presence of beta-
 CC amyloid plaques and as a dietary supplement for babies, small children,
 CC adults and senile persons, who have been subjected to chemotherapy or
 CC have suffered from cachexia or weight loss due to the chronic disease.
 CC Colostrinin peptides are also used as food additives and as an auxiliary
 CC withdrawal treatment for drug addicts, after a period of detoxification
 CC and in persons dependent on stimulants. Colostrinin peptides are used to
 CC prepare antibodies and to treat emotional disturbances, e.g. emotional
 CC disturbances of psychiatric patients in a state of depression. These
 CC colostrinin peptides improves the development of immune system in a new
 CC born child and to correct the immunological deficiencies in a child. The
 CC present sequence is colostrinin peptide 1 related to the invention
 XX Sequence 11 AA;
 SQ Query Match 62.9%; Score 56; DB 4; Length 11;
 Best Local Similarity 100.0%; Pred. No. 0.039;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 5 QPLLQVMEPEQ 15
 Db 1 QPLLQVMEPEQ 11
 RESULT 12
 AAEE07195
 ID AAEE07195 standard; peptide; 12 AA.
 XX AAEE07195;
 AC AAEE07195;
 XX 06-NOV-2001 (first entry)
 DT Colostrinin peptide 1.
 DE Colostrinin; nootropic; neuroprotective; immunomodulatory; antibacterial;
 XX Parkinson's disease; Alzheimer's disease; mental disorder; food additive;
 KW central nervous system disorder; neurodegenerative disorder; weight loss;
 KW beta-amyloid plaque; psychosis; neurosis; cachexia; motor neuron disease;
 KW acquired immunological deficiency; neurological disorder; dementia;
 KW antiviral.
 XX Unidentified.
 OS WO200155199-A1.
 XX 02-AUG-2001.
 XX 26-JAN-2001; 2001WO-GB000329.
 XX 26-JAN-2000; 2000GB-00001825.
 XX (REG-) REGEN THERAPEUTICS PLC.
 XX Georgiades JA;
 XX WPI; 2001-488775/53.
 XX Peptide useful as an interalia in the treatment of e.g. disorders of the
 PT immune system and the central nervous system comprises ten amino-terminal
 XX amino acid sequence derived from peptides present in colostrinin.
 PS Claim 1; Page 15; 40pp; English.
 XX The invention relates to colostrinin peptide fragments which are useful,
 CC inter alia, in the treatment of chronic disorders of the immune system
 CC and the central nervous system. Colostrinin peptides are used as a
 CC medicament in the treatment of neurological disorders e.g., dementia,
 CC neurodegenerative disorders e.g., Alzheimer's disease, motor neuron
 CC disease e.g., Parkinson's disease, mental disorders e.g., psychosis and
 CC neurosis, in acquired immunological deficiencies, chronic bacterial and
 CC viral infections and diseases characterised by the presence of beta-
 CC amyloid plaques and as a dietary supplement for babies, small children,
 CC adults and senile persons, who have been subjected to chemotherapy or
 CC have suffered from cachexia or weight loss due to the chronic disease.
 CC Colostrinin peptides are also used as food additives and as an auxiliary
 CC withdrawal treatment for drug addicts, after a period of detoxification
 CC and in persons dependent on stimulants. Colostrinin peptides are used to
 CC prepare antibodies and to treat emotional disturbances, e.g. emotional
 CC disturbances of psychiatric patients in a state of depression. These
 CC colostrinin peptides improves the development of immune system in a new
 CC born child and to correct the immunological deficiencies in a child. The
 CC present sequence is colostrinin peptide 1 related to the invention
 XX Sequence 11 AA;
 SQ Query Match 62.9%; Score 56; DB 4; Length 11;
 Best Local Similarity 100.0%; Pred. No. 0.039;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DE Modified colostrinin cyclic peptide #1.
 XX Colostrinin; nootropic; neuroprotective; immunomodulatory; antibacterial;
 KW Parkinson's disease; Alzheimer's disease; mental disorder; food additive;
 KW central nervous system disorder; neurodegenerative disorder; weight loss;
 KW beta-amyloid plaque; psychosis; neurosis; cachexia; motor neuron disease;
 KW acquired immunological deficiency; neurological disorder; dementia;
 KW antiviral; cyclic.
 XX Synthetic.
 XX Key Location/Qualifiers
 FT Modified-site 1
 FT /note= "N-terminal acetyl; this residue forms a cyclic
 XX linkage with Gln found at the C-terminal end"
 XX WO200155199-A1.
 XX 02-AUG-2001.
 XX 26-JAN-2001; 2001WO-GB000329.
 XX 26-JAN-2000; 2000GB-00001825.
 XX (REG-) REGEN THERAPEUTICS PLC.
 XX Georgiades JA;
 XX WPI; 2001-488775/53.
 XX Peptide useful as an interalia in the treatment of e.g. disorders of the
 PT immune system and the central nervous system comprises ten amino-terminal
 XX amino acid sequence derived from peptides present in colostrinin.
 PS Example 2; Page 8; 40pp; English.
 XX The invention relates to colostrinin peptide fragments which are useful,
 CC inter alia, in the treatment of chronic disorders of the immune system
 CC and the central nervous system. Colostrinin peptides are used as a
 CC medicament in the treatment of neurological disorders e.g., dementia,
 CC neurodegenerative disorders e.g., Alzheimer's disease, motor neuron
 CC disease e.g., Parkinson's disease, mental disorders e.g., psychosis and
 CC neurosis, in acquired immunological deficiencies, chronic bacterial and
 CC viral infections and diseases characterised by the presence of beta-
 CC amyloid plaques and as a dietary supplement for babies, small children,
 CC adults and senile persons, who have been subjected to chemotherapy or
 CC have suffered from cachexia or weight loss due to the chronic disease.
 CC Colostrinin peptides are also used as food additives and as an auxiliary
 CC withdrawal treatment for drug addicts, after a period of detoxification
 CC and in persons dependent on stimulants. Colostrinin peptides are used to
 CC prepare antibodies and to treat emotional disturbances, e.g. emotional
 CC disturbances of psychiatric patients in a state of depression. These
 CC colostrinin peptides improves the development of immune system in a new
 CC born child and to correct the immunological deficiencies in a child. The
 CC present sequence is modified colostrinin cyclic peptide #1 related to the
 CC invention
 XX Sequence 12 AA;
 SQ Query Match 62.9%; Score 56; DB 4; Length 12;
 Best Local Similarity 100.0%; Pred. No. 0.042;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 5 QPLLQVMEPEQ 15
 Db 2 QPLLQVMEPEQ 12
 RESULT 13
 ABO64584
 ID ABO64584 standard; protein; 82 AA.
 XX ABO64584;
 AC ABO64584;

XX 29-JUL-2004 (first entry)
XX Klebsiella pneumoniae polypeptide seqid 11101.
XX DE
XX Recombinant expression vector; transcription regulatory element;
XX Klebsiella pneumoniae protein; antibacterial; vaccine.
XX KW
XX Klebsiella pneumoniae.
XX OS
XX US6610836-B1.
XX PN
XX 26-AUG-2003.
XX PD
XX 27-JAN-2000; 2000US-00489039.
XX PF
XX 27-JAN-1999; 99US-0117747P.
XX PR
XX (GENO-) GENOME THERAPEUTICS CORP.
XX PA
XX Breton GL, Osborne M;
XX PI WPI: 2003-895346/82.
XX DR N-PSDB; ACH98135.
XX DX
XX New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for
XX PT preparing a vaccine composition against Klebsiella pneumoniae.
XX PS Disclosure; SEQ ID NO 11101; 932pp; English.
XX XX
XX The invention describes a new isolated nucleic acid encoding a Klebsiella
XX CC pneumoniae polypeptide. Also described are: a recombinant expression
XX CC vector comprising the nucleic acid, operably linked to a transcription
XX CC regulatory element; and a cell comprising the recombinant expression
XX CC vector. The nucleic acid is useful for preparing a vaccine composition
XX CC against Klebsiella pneumoniae. This is the amino acid sequence of a
XX CC Klebsiella pneumoniae polypeptide of the invention
XX SQ Sequence 82 AA;
XX
XX Query Match 57.3%; Score 51; DB 7; Length 82;
XX Best Local Similarity 69.2%; Pred. No. 2.1;
XX Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
XX
XX QY 1 LQTPQPLQVME 13
XX |||||:|:
XX 29 LQTPQPLIDAMLE 41
XX DB
XX
XX RESULT 14
XX ABP27487
XX ID ABP27487 standard; protein; 692 AA.
XX AC
XX ABP27487;
XX DT
XX 02-JUL-2002 (first entry)
XX XX
XX Streptococcus polypeptide SEQ ID NO 4150.
XX DE
XX Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
XX KW group A streptococcus; Streptococcus pyogenes; antibacterial;
XX KW antiinflammatory; infection; vaccine; meningitis; gene therapy.
XX XX
XX Streptococcus pyogenes.
XX OS
XX WO200234771-A2.
XX PN
XX 02-MAY-2002.
XX PD
XX 29-OCT-2001; 2001WO-GB004789.
XX PF
XX 27-OCT-2000; 2000GB-00026333.
XX PR
XX 24-NOV-2000; 2000GB-00028727.
XX PR
XX (CHIR-) CHIRON SPA.
XX PA

07-MAR-2001; 2001GB-00005640.
(CHIR-) CHIRON SPA.
(GENO-) INST GENOMIC RES.
Telford J, Masignani V, Margarit Y Rosi, Grandi G, Fraser C;
Tettelin H;
WPI: 2002-352536/38.
DR N-PSDB; ABN68118.
DR
DR New Streptococcus protein for the treatment or prevention of infection or
PT disease caused by Streptococcus bacteria, such as meningitis, and for
PT detecting a compound that binds to the protein.
XX
XX Claim 1; Page 3571; 4525pp; English.
XX PS
XX The invention relates to a protein (ABP254113-ABP30895) from group B
CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
CC the specification. The proteins have antibacterial and antiinflammatory
CC activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and
CC antibodies that bind (I) are used in the manufacture of medicaments for
CC the treatment or prevention of infection or disease caused by
CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
CC Nucleic acids encoding (I) are used to detect Streptococcus in a
CC biological sample. (I) is used to determine whether a compound binds to
CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be
CC used as a vaccine or diagnostic composition. The disease caused by
CC Streptococcus that is prevented or treated may be meningitis. Nucleic
CC acid encoding (I) may be used to recombinantly produce (I) and may be
CC used in gene therapy. Antibodies to (I) are used for affinity
CC chromatography, immunoassays, and distinguishing/identifying
CC Streptococcus proteins
XX SQ Sequence 692 AA;
XX
XX Query Match 53.9%; Score 48; DB 5; Length 692;
XX Best Local Similarity 40.0%; Pred. No. 60;
XX Matches 6; Conservative 8; Mismatches 1; Indels 0; Gaps 0
XX
XX QY 1 LQTPQPLQVMEPQ 15
XX ::|::|::|:
XX DB 400 IEVPEFVIQLMVEPK 414
XX
XX RESULT 15
XX ABP30759
XX ID ABP30759 standard; protein; 692 AA.
XX AC
XX ABP30759;
XX DT
XX 02-JUL-2002 (first entry)
XX XX
XX Streptococcus polypeptide SEQ ID NO 10694.
XX DE
XX Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
XX KW group A streptococcus; Streptococcus pyogenes; antibacterial;
XX KW antiinflammatory; infection; vaccine; meningitis; gene therapy.
XX XX
XX Streptococcus agalactiae.
XX OS
XX WO200234771-A2.
XX PN
XX 02-MAY-2002.
XX PD
XX 29-OCT-2001; 2001WO-GB004789.
XX PF
XX 27-OCT-2000; 2000GB-00026333.
XX PR
XX 24-NOV-2000; 2000GB-00028727.
XX PR
XX 07-MAR-2001; 2001GB-00005640.
XX XX
XX (CHIR-) CHIRON SPA.
XX PA

Search completed: April 12, 2005, 08:18:15
Job time : 97.2 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 12, 2005, 07:49:30 ; Search time 19.9364 Seconds
(without alignments)
63.654 Million cell updates/sec

Title: US-10-691-330-2

Perfect score: 89

Sequence: 1 LQTPQPLLQVMEPQGD 17

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5A COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	89	100.0	17	4	US-09-641-803-2
2	51	57.3	82	4	US-09-489-039A-11101
3	46	51.7	66	4	US-09-248-796A-26848
4	46	51.7	471	4	US-09-107-433-4845
5	46	51.7	693	4	US-09-583-110-4077
6	45	50.6	667	2	US-08-718-661-2
7	43	48.3	351	4	US-09-248-796A-25420
8	42	47.2	141	4	US-09-270-767-35942
9	42	47.2	141	4	US-09-270-767-51159
10	42	47.2	230	4	US-09-248-796A-18983
11	42	47.2	328	4	US-09-248-796A-22224
12	41	46.1	99	4	US-09-493-565-2
13	41	46.1	294	4	US-09-328-352-7430
14	41	46.1	343	3	US-08-858-003-32
15	41	46.1	343	3	US-09-078-166-32
16	41	46.1	343	3	US-08-997-467-32
17	41	46.1	428	3	US-09-347-833-6
18	40.5	45.5	369	2	US-08-596-291-4
19	40.5	45.5	369	3	US-09-100-804-4
20	40.5	45.5	369	5	PCT-US94-09943-4
21	40.5	45.5	534	4	US-09-949-016-7882
22	40	44.9	85	4	US-09-621-976-5318
23	40	44.9	109	4	US-09-248-796A-26851
24	40	44.9	124	4	US-09-270-767-31669
25	40	44.9	124	4	US-09-270-767-46886
26	40	44.9	671	3	US-09-121-321-16
27	40	44.9	671	3	US-08-933-803A-16

Sequence 22785, A
Sequence 7812, Ap
Sequence 23561, A
Sequence 42966, A
Sequence 15, Appl
Sequence 30333, A
Sequence 6328, Ap
Sequence 456, App
Sequence 4213, Ap
Sequence 3608, Ap
Sequence 31345, A
Sequence 57079, A
Sequence 38, Appl
Sequence 8491, Ap
Sequence 41835, A
Sequence 43648, A
Sequence 10892, A
Sequence 6366, Ap

ALIGNMENTS

RESULT 1

US-09-641-803-2

; Sequence 2, Application US/09641803

; Patent No. 6500798

; GENERAL INFORMATION:

; APPLICANT: STANTON, G. John

; APPLICANT: HUGHES, Thomas K.

; APPLICANT: BOLDOGH, Istvan

; TITLE OF INVENTION: USE OF COLOSTRININ, CONSTITUENT PEPTIDES THEREOF, AND ANALOGS THEREOF AS OXIDATIVE STRESS REGULATORS

; FILE REFERENCE: 265.00220101

; CURRENT APPLICATION NUMBER: US/09/641.803

; PRIOR FILING DATE: 2000-08-17

; PRIOR APPLICATION NUMBER: 60/149,310

; PRIOR FILING DATE: 1999-08-17

; NUMBER OF SEQ ID NOS: 34

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 2

; LENGTH: 17

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: synthetic

; OTHER INFORMATION: peptide

US-09-641-803-2

Query Match 100.0%; Score 89; DB 4; Length 17;

Best Local Similarity 100.0%; Pred. No. 9.8e-08; Mismatches 0; Indels 0; Gaps 0;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LQTPQPLLQVMEPQGD 17

Db 1 LQTPQPLLQVMEPQGD 17

RESULT 2

US-09-489-039A-11101

; Sequence 11101, Application US/09489039A

; Patent No. 6610836

; GENERAL INFORMATION:

; APPLICANT: Gary Breton et. al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 2709.2004001

; CURRENT APPLICATION NUMBER: US/09/489.039A

; CURRENT FILING DATE: 2000-01-27

; PRIOR APPLICATION NUMBER: US 60/117,747

; PRIOR FILING DATE: 1999-01-29

; NUMBER OF SEQ ID NOS: 14342

; SEQ ID NO 11101

```
; LENGTH: 82
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-11101

Query Match          57.3%; Score 51; DB 4; Length 82;
Best Local Similarity 69.2%; Pred. No. 0.54;
Matches 9; Conservative 2; Mismatches 2; Indels 2; Gaps 0;

QY 1 LQTPQLQVMMPEQ 13
Db 29 LQTPQLQVMMPEQ 41

RESULT 3
US-09-248-796A-26848
; Sequence 26848, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 26848
; LENGTH: 66
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-26848

Query Match          51.7%; Score 46; DB 4; Length 66;
Best Local Similarity 41.2%; Pred. No. 2.6;
Matches 7; Conservative 5; Mismatches 5; Indels 5; Gaps 0;

QY 1 LQTPQLQVMMPEQGD 17
Db 48 LDAPEKVPVLEPKGD 64

RESULT 4
US-09-107-433-4845
; Sequence 4845, Application US/09107433
; Patent No. 6800744
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
; SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE FOR DIAGNO
; THERAPEUTICS
; NUMBER OF SEQUENCES: 5206
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: <Unknown>
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: <Unknown>
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,433
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/ 085131
; FILING DATE: May 12, 1998
```

```
; APPLICATION NUMBER: 60/051553
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-011
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 4845:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 471 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (B) LOCATION 1...471
; SEQUENCE DESCRIPTION: SEQ ID NO: 4845:
US-09-107-433-4845

Query Match          51.7%; Score 46; DB 4; Length 471;
Best Local Similarity 40.0%; Pred. No. 23;
Matches 6; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY 1 LQTPQLQVMMPEQ 15
Db 178 INVPEVQLMVEPK 192

RESULT 5
US-09-583-110-4077
; Sequence 4077, Application US/09583110
; Patent No. 6699703
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; FILE REFERENCE: PATH00-07A
; CURRENT APPLICATION NUMBER: US/09/583,110
; CURRENT FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 4077
; LENGTH: 693
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-4077

Query Match          51.7%; Score 46; DB 4; Length 693;
Best Local Similarity 40.0%; Pred. No. 35;
Matches 6; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY 1 LQTPQLQVMMPEQ 15
Db 400 INVPEVQLMVEPK 414

RESULT 6
US-08-718-661-2
; Sequence 2, Application US/08718661
; Patent No. 5876972
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Nucleic acid molecules coding for mammalian
```


;; TITLE OF INVENTION: tumor suppressor proteins and methods for their isolation
;; NUMBER OF SEQUENCES: 15
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk

;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS

;; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
;; CURRENT APPLICATION DATA:

;; APPLICATION NUMBER: US/08/718,661

;; FILING DATE:

;; CLASSIFICATION: 530

;; INFORMATION FOR SEQ ID NO: 2:

;; SEQUENCE CHARACTERISTICS:

;; LENGTH: 667 amino acids

;; TYPE: amino acid

;; TOPOLOGY: linear

;; MOLECULE TYPE: protein

US-08-718-661-2

Query Match 50.6%; Score 45; DB 2; Length 667;
Best Local Similarity 60.0%; Pred. No. 48;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 1 LQTPQLLQVMEPQ 15

Db 543 LPVQPLPQPMQ 557

RESULT 7

US-09-248-796A-25420

;; Sequence 25420, Application US/09248796A

;; Patent No. 6747137

;; GENERAL INFORMATION:

;; APPLICANT: Keith Weinstock et al

;; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN

;; FILE REFERENCE: 107196.132

;; CURRENT APPLICATION NUMBER: US/09/248,796A

;; PRIOR FILING DATE: 1999-02-12

;; PRIOR APPLICATION NUMBER: US 60/074,725

;; PRIOR FILING DATE: 1998-02-13

;; PRIOR APPLICATION NUMBER: US 60/096,409

;; PRIOR FILING DATE: 1998-08-13

;; NUMBER OF SEQ ID NOS: 28208

;; SEQ ID NO 25420

;; LENGTH: 351

;; TYPE: PRT

;; ORGANISM: Candida albicans

US-09-248-796A-25420

Query Match 48.3%; Score 43; DB 4; Length 351;
Best Local Similarity 63.6%; Pred. No. 49;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 4 PQLLQVMEP 14

Db 9 PQLLQPLLP 19

RESULT 8

US-09-270-767-35942

;; Sequence 35942, Application US/09270767

;; Patent No. 6703491

;; GENERAL INFORMATION:

;; APPLICANT: Homburger et al.

;; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster

;; FILE REFERENCE: File Reference: 7326-094

;; CURRENT APPLICATION NUMBER: US/09/270,767

;; CURRENT FILING DATE: 1999-03-17

;; NUMBER OF SEQ ID NOS: 62517

;; SOFTWARE: PatentIn Ver. 2.0

;; SEQ ID NO 35942

;; LENGTH: 141

;; TYPE: PRT
;; ORGANISM: Drosophila melanogaster
US-09-270-767-35942

Query Match 47.2%; Score 42; DB 4; Length 141;
Best Local Similarity 52.9%; Pred. No. 26;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 LQTPQLLQVMEPQ 17

Db 7 LQTTMLVQVALQ 23

RESULT 9

US-09-270-767-51159

;; Sequence 51159, Application US/09270767

;; Patent No. 6703491

;; GENERAL INFORMATION:

;; APPLICANT: Homburger et al.

;; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster

;; FILE REFERENCE: File Reference: 7326-094

;; CURRENT APPLICATION NUMBER: US/09/270,767

;; CURRENT FILING DATE: 1999-03-17

;; NUMBER OF SEQ ID NOS: 62517

;; SOFTWARE: PatentIn Ver. 2.0

;; SEQ ID NO 51159

;; LENGTH: 141

;; TYPE: PRT

;; ORGANISM: Drosophila melanogaster

US-09-270-767-51159

Query Match 47.2%; Score 42; DB 4; Length 141;
Best Local Similarity 52.9%; Pred. No. 26;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 LQTPQLLQVMEPQ 17

Db 7 LQTTMLVQVALQ 23

RESULT 10

US-09-248-796A-18983

;; Sequence 18983, Application US/09248796A

;; Patent No. 6747137

;; GENERAL INFORMATION:

;; APPLICANT: Keith Weinstock et al

;; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN

;; FILE REFERENCE: 107196.132

;; CURRENT APPLICATION NUMBER: US/09/248,796A

;; CURRENT FILING DATE: 1999-02-12

;; PRIOR APPLICATION NUMBER: US 60/074,725

;; PRIOR FILING DATE: 1998-02-13

;; PRIOR APPLICATION NUMBER: US 60/096,409

;; PRIOR FILING DATE: 1998-08-13

;; NUMBER OF SEQ ID NOS: 28208

;; SEQ ID NO 18983

;; LENGTH: 230

;; TYPE: PRT

;; ORGANISM: Candida albicans

US-09-248-796A-18983

Query Match 47.2%; Score 42; DB 4; Length 230;
Best Local Similarity 53.3%; Pred. No. 44;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 2 QTPQLLQVMEPQ 16

Db 157 QPFPNGMMPPG 171

RESULT 11

US-09-248-796A-22224

```

; ORGANISM: Acinetobacter baumannii
US-09-328-352-7430

Query Match          46.1%; Score 41; DB 4; Length 294;
Best Local Similarity 37.9%; Pred. No. 83;
Matches 11; Conservative 4; Mismatches 2; Indels 1

Qy 1 LQTPQ-----PLL-----QVMMPEQGD 17
   |||||
Db 151 LTPQDIPILSPILIKEIFYRLILMSQGD 179
   |||:|:|:|:|

RESULT 14
US-08-858-003-32
; Sequence 32, Application US/08858003
; Patent No. 6060234
; GENERAL INFORMATION:
; APPLICANT: Katz, Leonard
; APPLICANT: Stassi, Diane L.
; APPLICANT: Summers Jr., Richard G.
; APPLICANT: Ruan, Xiaocan
; APPLICANT: Pereda-Lopez, Ana
; APPLICANT: Kakavas, Stephan J.
; TITLE OF INVENTION: NOVEL POLYKETIDE DERIVATIVES
; TITLE OF INVENTION: NOVEL POLYKETIDE DERIVATIVES
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Abbott Laboratories
; STREET: 100 Abbott Park Rd.
; CITY: Abbott Park
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/858,003
; FILING DATE: 16-MAY-1979
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Dianne Casuto
; REGISTRATION NUMBER: P-40,943
; REFERENCE/DOCKET NUMBER: 4952.US.P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (847)-938-3137
; TELEFAX: (847)-938-2623
; TELEX:
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 343 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: No. 6060234e
US-08-858-003-32

Query Match          46.1%; Score 41; DB 3; Length 343;
Best Local Similarity 46.7%; Pred. No. 99;
Matches 7; Conservative 3; Mismatches 5; Indels 0

Qy 2 QTPQPLQVMMPEQ 16
   |||||
Db 42 QLPRLKDVLFAPEG 56
   |||:|:|:|:|

RESULT 15
US-09-078-166-32

```

; Sequence 32, Application US/09078166
; Patent No. 6063561
; GENERAL INFORMATION:
; APPLICANT: Katz, Leonard
; APPLICANT: Stassi, Diane L.
; APPLICANT: Summers Jr., Richard G.
; APPLICANT: Ruan, Xiaolan
; APPLICANT: Pereda-Lopez, Ana
; APPLICANT: Kakavas, Stephan J.
; TITLE OF INVENTION: NOVEL POLYKETIDE DERIVATIVES
; TITLE OF INVENTION: AND RECOMBINANT METHODS FOR MAKING SAME
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Abbott Laboratories
; STREET: 100 Abbott Park Rd.
; CITY: Abbott Park
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/078,166
; FILING DATE: 16-MAY-1979
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Dianne Casuto
; REGISTRATION NUMBER: P-40,943
; REFERENCE/DOCKET NUMBER: 4952.US.P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (847)-938-3137
; TELEFAX: (847)-938-2623
; TELEX:
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 343 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: No. 6063561e
; US-09-078-166-32

Query Match 46.1%; Score 41; DB 3; Length 343;
Best Local Similarity 46.7%; Pred. No. 99;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
Qy 2 QTPQPLQVMEPQG 16
Db 42 QLRPLKDVLEAPEG 56

Search completed: April 12, 2005, 08:07:49
Job time : 20.9364 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 12, 2005, 07:49:30 ; Search time 52.7 Seconds
(without alignments)
107.096 Million cell updates/sec

Title: US-10-691-330-2

Perfect score: 89

Sequence: 1 LQTPQPLLQVMEPQGD 17

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1418010 seqs, 331997259 residues

Total number of hits satisfying chosen parameters: 1418010

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	89	100.0	17	14	US-10-281-652-2
2	89	100.0	17	17	US-10-691-330-2
3	56	62.9	11	16	US-10-182-110-1
4	48	53.9	692	15	US-10-282-122A-74378
5	48	53.9	693	15	US-10-282-122A-72253
6	46	51.7	693	9	US-09-815-242-13233
7	46	51.7	693	15	US-10-282-122A-73764
8	46	51.7	693	17	US-10-472-928-374
9	45	50.6	74	16	US-10-767-701-49828
10	45	50.6	83	9	US-09-864-761-35253
11	44	49.4	50	14	US-10-223-978-2
12	44	49.4	134	17	US-10-826-157-4
13	44	49.4	549	15	US-10-425-114-51657

Sequence 2, Appli
Sequence 2, Appli
Sequence 1, Appli
Sequence 74378, A
Sequence 72253, A
Sequence 13233, A
Sequence 73764, A
Sequence 374, App
Sequence 49828, A
Sequence 35253, A
Sequence 2, Appli
Sequence 4, Appli
Sequence 51657, A

Query Match 100.0%; Score 89; DB 14; Length 17;

Best Local Similarity 100.0%; Pred. No. 4.5e-07;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LQTPQPLLQVMEPQGD 17

Db 1 LQTPQPLLQVMEPQGD 17

ALIGNMENTS

RESULT 1

US-10-281-652-2
; Sequence 2, Application US/10281652
; Publication No. US20030091608A1
; GENERAL INFORMATION:
; APPLICANT: STANTON, G. John
; APPLICANT: HUGHES, Thomas K.
; APPLICANT: BOLDGCH, Istvan
; TITLE OF INVENTION: USE OF COLOSTRININ, CONSTITUENT PEPTIDES THEREOF, AND
; FILE REFERENCE: 265.00220101
; CURRENT APPLICATION NUMBER: US/10/281.652
; CURRENT FILING DATE: 2002-10-28
; PRIOR FILING DATE: 2000-08-17
; PRIOR FILING DATE: 60/149,310
; PRIOR FILING DATE: 1999-08-17
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: peptide
US-10-281-652-2

Sequence 10, Appli
Sequence 824, App
Sequence 126258,
Sequence 10, Appli
Sequence 68, Appli
Sequence 94, Appli
Sequence 94, Appli
Sequence 90, Appli
Sequence 595, App
Sequence 10, Appli
Sequence 10, Appli
Sequence 10, Appli
Sequence 10, Appli
Sequence 10, Appli
Sequence 10, Appli
Sequence 10, Appli
Sequence 2, Appli
Sequence 2, Appli
Sequence 1058, Ap
Sequence 900, App
Sequence 10, Appli
Sequence 92, Appli
Sequence 92, Appli
Sequence 57461, A
Sequence 61020, A
Sequence 804, App
Sequence 2260, Ap
Sequence 3898, Ap
Sequence 46468, A

```

RESULT 2
US-10-691-330-2
; Sequence 2, Application US/10691330
; Publication No. US20050042300A1
; GENERAL INFORMATION:
; APPLICANT: Boldog, Istvan
; APPLICANT: Stanton, G. John
; APPLICANT: Georgiades, Jerzy A.
; APPLICANT: Hughes, Thomas K., Jr.
; APPLICANT: Kruzel, Marian
; TITLE OF INVENTION: USE OF COLOSTRININ, CONSTITUENT PEPTIDES THEREOF, AND ANALOGS
; FILE REFERENCE: 265, 00390101
; CURRENT APPLICATION NUMBER: US/10/691,330
; CURRENT FILING DATE: 2003-10-22
; PRIOR APPLICATION NUMBER: US 60/420,369
; PRIOR FILING DATE: 2002-10-22
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
; LENGTH: 17
; TYPE: PRT
; ORGANISM: artificial
; FEATURE:
; OTHER INFORMATION: Synthetic Peptides
US-10-691-330-2

Query Match 100.0%; Score 89; DB 17; Length 17;
Best Local Similarity 100.0%; Pred. No. 4.5e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LQTPQLLVQVMPEQD 17
| | | | | | | | | | | | | | | |
Db 1 LQTPQLLVQVMPEQD 17

RESULT 3
US-10-182-110-1
; Sequence 1, Application US/10182110
; Publication No. US20040171553A1
; GENERAL INFORMATION:
; APPLICANT: Regen Therapeutics plc
; APPLICANT: Georgiades, Jerzy A.
; TITLE OF INVENTION: Peptide Fragments of Colostrinin And Their Use
; FILE REFERENCE: AT-14173
; CURRENT APPLICATION NUMBER: US/10/182,110
; CURRENT FILING DATE: 2003-04-21
; PRIOR APPLICATION NUMBER: GB0001825.9
; PRIOR FILING DATE: 2000-01-26
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Truncated version of a peptide found in colostrinin
US-10-182-110-1

Query Match 62.9%; Score 56; DB 16; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.039;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 QPLLVQVMPEQ 15
| | | | | | | | | | |
Db 1 QPLLVQVMPEQ 11

RESULT 4
US-10-282-122A-74378
; Sequence 74378, Application US/10282122A
; Publication No. US20040029129A1

```

```

; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining prior application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 74378
; LENGTH: 692
; TYPE: PRT
; ORGANISM: Streptococcus pyogenes
US-10-282-122A-74378

Query Match 53.9%; Score 48; DB 15; Length 692;
Best Local Similarity 40.0%; Pred. No. 59;
Matches 6; Conservative 8; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LQTPQLLVQVMPEQ 15
| | | | | | | | | | |
Db 400 IEVPEPVIQLMVEPK 414

RESULT 5
US-10-282-122A-72253
; Sequence 72253, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

```

```

; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13233
; LENGTH: 693
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-815-242-13233

Query Match          51.7%; Score 46; DB 9; Length 693;
Best Local Similarity 40.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY      1 LQTQPQLQVMEPQ 15
       : |::|::|::|:
Db      400 INVPEPVIQLMVEPK 414

RESULT 7
US-10-282-122A-73764
; Sequence 73764, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 73764
; LENGTH: 693
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-282-122A-73764

Query Match          51.7%; Score 46; DB 15; Length 693;
Best Local Similarity 40.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY      1 LQTQPQLQVMEPQ 15
       : |::|::|::|:

```

GENERAL INFORMATION:

1 LOTPOPLLOVMMEPO 15


```
Db      21 LKQPQLLASLPQ 35

RESULT 11
US-10-223-978-2
; Sequence 2, Application US/10223978
; Publication No. US20030125522A1
; GENERAL INFORMATION:
; APPLICANT: Atgen Co., Ltd.
; APPLICANT: Kim, Jong-Sun
; TITLE OF INVENTION: No. US20030125522A1el Peptides Conferring Environmental Stress Re
; TITLE OF INVENTION: Proteins Including Said Peptides
; FILE REFERENCE: 59520-00003
; CURRENT APPLICATION NUMBER: US/10/223,978
; CURRENT FILING DATE: 2002-08-20
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 2
; LENGTH: 50
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(50)
; OTHER INFORMATION: Acidic tail amino acid sequence 85-134 of beta-synuclein
US-10-223-978-2

Query Match      49.4%; Score 44; DB 14; Length 50;
Best Local Similarity 46.2%; Pred. No. 15;
Matches 6; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

Qy      5 OPLQVMMPEQD 17
      :||:::||::|:
Db      21 EPLIEPLMEPEGE 33

RESULT 12
US-10-826-157-4
; Sequence 4, Application US/10826157
; Publication No. US20050064548A1
; GENERAL INFORMATION:
; APPLICANT: Lindquist, Susan L.
; APPLICANT: Outeiro, Tiago
; TITLE OF INVENTION: YEAST ECOTOPICALLY EXPRESSING ABNORMALLY
; TITLE OF INVENTION: PROCESSED PROTEINS AND USES THEREFOR
; FILE REFERENCE: 17481-003001
; CURRENT APPLICATION NUMBER: US/10/826,157
; CURRENT FILING DATE: 2004-04-16
; PRIOR APPLICATION NUMBER: US 60/472,317
; PRIOR FILING DATE: 2003-05-20
; PRIOR APPLICATION NUMBER: US 60/463,284
; PRIOR FILING DATE: 2003-04-16
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 134
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-826-157-4

Query Match      49.4%; Score 44; DB 17; Length 134;
Best Local Similarity 46.2%; Pred. No. 43;
Matches 6; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

Qy      5 OPLQVMMPEQD 17
      :||:::||::|:
Db      105 EPLIEPLMEPEGE 117

RESULT 13
US-10-425-114-51657
; Sequence 51657, Application US/10425114
; Publication No. US20040034888A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 51657
; LENGTH: 549
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3150-095-C11_FLI.pep
US-10-425-114-51657

Query Match      49.4%; Score 44; DB 15; Length 549;
Best Local Similarity 56.2%; Pred. No. 1.9e+02;
Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy      1 LQTPQLQVMMPEQ 16
      |:::||::|:
Db      324 LRYPOPLLSVAVSPG 339

RESULT 14
US-09-789-054A-10
; Sequence 10, Application US/09789054A
; Publication No. US20020184659A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Zhu, Qun
; TITLE OF INVENTION: PLANT GENES ENCODING DR1 AND DRAP1, A GLOBAL REPRESSOR COMPLEX OF
; TITLE OF INVENTION: TRANSCRIPTION
; FILE REFERENCE: BB107 US CIP
; CURRENT APPLICATION NUMBER: US/09/789,054A
; CURRENT FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 09/485558
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: PCT/US98/16688
; PRIOR FILING DATE: 1998-08-12
; PRIOR APPLICATION NUMBER: 60/055,865
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 10
; LENGTH: 296
; TYPE: PRT
; ORGANISM: Oryza sativa
US-09-789-054A-10

Query Match      48.3%; Score 43; DB 9; Length 296;
Best Local Similarity 64.3%; Pred. No. 1.4e+02;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy      2 QTPQPLQVMMPEQ 15
      |||||:|:
Db      199 QTPQPPQVHPQPQ 212

RESULT 15:
US-10-374-780A-824
; Sequence 824, Application US/10374780A
; Publication No. US20040019927A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, Bradley K
; APPLICANT: Riechmann, Jose Luis
; APPLICANT: Jiang, Cai-Zhong
```

APPLICANT: Heard, Jacqueline E
APPLICANT: Haake, Volker
APPLICANT: Creelman, Robert A
APPLICANT: Ratcliffe, Oliver
APPLICANT: Adam, Luc J
APPLICANT: Reuber, T. Lynne
APPLICANT: Keddle, James
APPLICANT: Broun, Pierre E
APPLICANT: Pilgrim, Marsha L
APPLICANT: Dubeil III, Arnold T
APPLICANT: Pineda, Omaira
APPLICANT: Yu, Guo-Liang
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES IN PLANTS
FILE REFERENCE: MBI-0047 CIP
CURRENT APPLICATION NUMBER: US/10/374,780A
CURRENT FILING DATE: 2003-02-25
PRIOR APPLICATION NUMBER: 09/837,944
PRIOR FILING DATE: 2001-04-18
PRIOR APPLICATION NUMBER: 60/310,847
PRIOR FILING DATE: 2001-08-09
PRIOR APPLICATION NUMBER: 09/934,455
PRIOR FILING DATE: 2001-08-22
PRIOR APPLICATION NUMBER: 60/336,049
PRIOR FILING DATE: 2001-11-19
PRIOR APPLICATION NUMBER: 60/338,692
PRIOR FILING DATE: 2001-12-11
PRIOR APPLICATION NUMBER: 10/171,468
PRIOR FILING DATE: 2002-06-14
PRIOR APPLICATION NUMBER: 10/225,066
PRIOR FILING DATE: 2002-08-09
PRIOR APPLICATION NUMBER: 10/225,067
PRIOR FILING DATE: 2002-08-09
PRIOR APPLICATION NUMBER: 10/225,068
PRIOR FILING DATE: 2002-08-09
NUMBER OF SEQ ID NOS: 2906
SOFTWARE: PatentIn version 3.2
SEQ ID NO 824
LENGTH: 296
TYPE: PRT
ORGANISM: Oryza sativa
FEATURE:
OTHER INFORMATION: Orthologous to G484.
US-10-374-780A-824

Query Match 48.3%; Score 43; DB 15; Length 296;
Best Local Similarity 64.3%; Pred. No. 1.4e+02;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 2 QTPQPLQVMMEPQ 15
||| | | |
Db 199 QTPQFPQVHPQPQ 212

Search completed: April 12, 2005, 07:55:20
Job time : 52.7 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 12, 2005, 07:49:31 ; Search time 12.2091 Seconds
(without alignments)
133.973 Million cell updates/sec

Title: US-10-691-330-2

Perfect score: 89

Sequence: 1 LQTPQPLLQVMBEPQGD 17

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_79.*

1: Pirl.*

2: Pirl.*

3: Pirl.*

4: Pirl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48	53.9	709	2 D86907	elongation factor
2	47.5	53.4	192	2 A11753	Orf50 [bacteriophage
3	46	51.7	639	2 S03268	tetracycline resist
4	46	51.7	693	2 B97903	elongation factor
5	46	51.7	693	2 B95032	translation elonga
6	44	49.4	134	2 S44430	synuclein - human
7	44	49.4	134	2 S39046	phosphonuroprotei
8	44	49.4	137	2 I56498	phosphonuroprotei
9	44	49.4	671	2 C96534	probable Poly-A Bi
10	43	48.3	134	2 A80784	conserved hypothet
11	43	48.3	639	2 A60633	tetracycline resist
12	43	48.3	639	2 A56779	tetracycline resist
13	43	48.3	639	2 S13142	tetracycline resist
14	43	48.3	695	2 A51406	translation elonga
15	43	48.3	695	2 A51782	translation elonga
16	43	48.3	699	2 E37594	elongation factor
17	43	48.3	699	2 AC2816	translation elonga
18	43	48.3	844	2 F86231	hypothetical prote
19	42	47.2	263	2 S74353	carbonic anhydrase
20	42	47.2	580	2 S13328	hypothetical prote
21	42	47.2	595	2 F85438	nucleoporin-like p
22	42	47.2	643	2 S55610	polyprotein - equi
23	42	47.2	698	2 C82332	translation elonga
24	41	46.1	78	2 D81160	hypothetical prote
25	41	46.1	78	2 C86006	hypothetical prote
26	41	46.1	639	2 E65136	hypothetical prote
27	41	46.1	639	2 J24333	tetracycline resist
28	41	46.1	641	2 JN0800	tetracycline-minoc
29	41	46.1	857	2 JC4169	phosphoenolpyruvat

UL52 protein - hum
hypothetical prote
hypothetical prote
hypothetical prote
protein-tyrosine-p
protein-tyrosine-p
protein-tyrosine-p
hypothetical prote
hypothetical prote
lactoylglutathione
hypothetical prote
protein FIN1.10 (e
esterase, probable
translation elonga

ALIGNMENTS

RESULT 1

D86907

elongation factor G [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
C;Species: Lactococcus lactis subsp. lactis
C;Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 09-Jul-2004
C;Accession: D86907

R;Bolotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarne, K.; Weissenbach, J.; Ehrlich

Genome Res. 11, 731-753, 2001

A;Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis se

A;Reference number: A86625; MUID:21235186; PMID:11337471

A;Accession: D86907

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-709 <STO>

A;Cross-references: UNIPROT:Q9CDG1; GB:AE005176; PID:g12725331; PIDN:AAK06358.1; GSPDB:G

A;Experimental source: strain IL1403

C;Genetics:

A;Gene: fusA

C;Superfamily: translation elongation factor G; translation elongation factor Tu homolog

Query Match 53.9%; Score 48; DB 2; Length 709;

Best Local Similarity 40.0%; Pred. No. 5.8;

Matches 6; Conservative 8; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LQTPQPLLQVMBEPQ 15

Db 415 IEVPEPVIQLMVEPK 429

RESULT 2

AP1753

Orf50 [bacteriophage b11285] homolog lin2571 [imported] - Listeria innocua (strain Clp1

C;Species: Listeria innocua

C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004

C;Accession: AF1753

R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker

; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.

D.; Jones, L.M.; Karst, U.

Science 294, 849-852, 2001

A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma

ok, C.; Schluster, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,

A;Title: Comparative genomics of Listeria species.

A;Reference number: AB1077; MUID:21537279; PMID:11679669

A;Accession: AF1753

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-192 <GLA>

A;Cross-references: UNIPROT:Q928G4; GB:AL592022; PIDN:CAC97798.1; PID:g16415093; GSPDB:G

A;Experimental source: strain Clp11262

C;Genetics:

A;Gene: lin2571

Query Match 53.4%; Score 47.5; DB 2; Length 192;
 Best Local Similarity 50.0%; Pred. No. 1.6;
 Matches 9; Conservative 5; Mismatches 1; Indels 3; Gaps 1;

QY 3 TPQPL---LQVMMPEQGD 17
 |||:| :|:|:|:|
 Db 31 TPKPLPGGVISVPEQGD 48

RESULT 3

S03268
 tetraacycline resistance protein tetM - Ureaplasma urealyticum
 C;Species: Ureaplasma urealyticum
 C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 09-Jul-2004
 C;Accession: S03268
 R;Sanchez-Pescador, R.; Brown, J.T.; Roberts, M.; Urdea, M.S.
 Nucleic Acids Res. 16, 1216-1217, 1988
 A;Title: The nucleotide sequence of the tetracycline resistance determinant tetM from *U. urealyticum*
 A;Reference number: S03268; MUID:88144009; PMID:3344217
 A;Accession: S03268
 A;Molecule type: DNA
 A;Residues: 1-639 <S>
 A;Cross-references: UNIPROT:P09757; GB:U08812; EMBL:X06901; NID:9475983; PIDN:AAA73978.1
 C;Genetics:
 A;Gene: tetM
 A;Genetic code: SGC3
 C;Superfamily: translation elongation factor G; translation elongation factor Tu homolog

C;Keywords: antibiotic resistance; GTP binding; nucleotide binding; P-loop
 F;4-131/Domain: translation elongation factor Tu homolog <ETU>
 F;10-17/Region: nucleotide-binding motif A (P-loop)
 F;128-131/Region: nucleotide-binding motif A (P-loop)
 F;220-222/Region: GTP-binding NKXD motif
 F;16,17,55,128,129,131,220/Binding site: Mg-GTP (Lys, Thr, Asn, Lys, Asp, Ser) #sta

Query Match 51.7%; Score 46; DB 2; Length 639;
 Best Local Similarity 50.0%; Pred. No. 11;
 Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 LQTPQLQVMMPEQ 14
 :|:|:|:|:|
 Db 339 IENPHLLQIVPEP 352

RESULT 4

B97903
 elongation factor G [imported] - Streptococcus pneumoniae (strain R6)
 C;Species: Streptococcus pneumoniae
 C;Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
 C;Accession: B97903
 R;Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; H
 e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M
 y, P.; Sun, P.M.; Winkler, M.E.
 J. Bacteriol. 183, 5709-5717, 2001
 A;Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;
 A;Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
 A;Reference number: A97872; MUID:21429245; PMID:11544234
 A;Accession: B97903
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-693 <KUR>
 A;Cross-references: UNIPROT:Q97SQ3; GB:AB007317; PIDN:AAK99054.1; PID:gl5457799; GSPDB:G
 C;Genetics:
 A;Gene: fusA
 C;Superfamily: translation elongation factor G; translation elongation factor Tu homolog

Query Match 51.7%; Score 46; DB 2; Length 693;
 Best Local Similarity 40.0%; Pred. No. 12;
 Matches 6; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY 1 LQTPQLQVMMPEQ 15
 :|:|:|:|:|
 Db 400 INVPEPQLMVEPK 414

RESULT 5

B95032
 translation elongation factor G [imported] - Streptococcus pneumoniae (strain TIGR4)
 C;Species: Streptococcus pneumoniae
 C;Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 09-Jul-2004
 C;Accession: B95032
 R;Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid
 on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple, I.
 nson, T.; Hickey, E.K.; Holt, I.E.
 Science 293, 498-506, 2001
 A;Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
 A;Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
 A;Reference number: A95000; MUID:21357209; PMID:11463916
 A;Accession: B95032
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-693 <KUR>

A;Cross-references: UNIPROT:Q97SQ3; GB:AB005672; PIDN:AAK74451.1; PID:gl4971744; GSPDB:GN
 A;Experimental source: strain TIGR4
 C;Genetics:
 A;Gene: SP0273
 C;Superfamily: translation elongation factor G; translation elongation factor Tu homolog

Query Match 51.7%; Score 46; DB 2; Length 693;
 Best Local Similarity 40.0%; Pred. No. 12;
 Matches 6; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY 1 LQTPQLQVMMPEQ 15
 :|:|:|:|:|
 Db 400 INVPEPQLMVEPK 414

RESULT 6

S44430
 synuclein - human
 C;Species: Homo sapiens (man)
 C;Date: 07-Oct-1994 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
 C;Accession: S44430
 R;Jakes, R.; Spillantini, M.G.; Goedert, M.
 FEBS Lett. 345, 27-32, 1994
 A;Title: Identification of two distinct synucleins from human brain.
 A;Reference number: S44430; MUID:94252398; PMID:8194594
 A;Accession: S44430
 A;Status: preliminary
 A;Molecule type: mRNA
 A;Residues: 1-134 <JAK>
 A;Cross-references: UNIPROT:Q16143; GB:S69965; NID:G546911; PIDN:AAB30860.1; PID:G546912

Query Match 49.4%; Score 44; DB 2; Length 134;
 Best Local Similarity 46.2%; Pred. No. 4.2;
 Matches 6; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

QY 5 QPLQVMMPEQGD 17
 :|:|:|:|:|
 Db 105 EPLIEPLMEPEGR 117

RESULT 7

S39046
 phosphonoprotein 14 - bovine
 N;Alternate names: brain-specific protein, 14K
 C;Species: Bos primigenius taurus (cattle)
 C;Date: 19-May-1994 #sequence_revision 05-May-1995 #text_change 09-Jul-2004
 C;Accession: S39046; J00238
 R;Nakajo, S.; Teukada, K.; Omata, K.; Nakamura, Y.; Nakaya, K.
 Eur. J. Biochem. 217, 1057-1063, 1993
 A;Title: A new brain-specific 14-kDa protein is a phosphoprotein. Its complete amino acid
 A;Reference number: S39046; MUID:94039126; PMID:8223629
 A;Accession: S39046
 A;Status: preliminary
 A;Molecule type: protein
 A;Residues: 1-134 <NAK>

A;Cross-references: UNIPROT:P33567
 R;Nakajo, S.; Tsukada, K.; Omata, K.; Nakamura, Y.; Nakaya, K.
 submitted to JIPID, October 1993
 A;Description: A new brain-specific 14-kDa protein is a phosphoprotein: Its complete amino acid sequence is deduced from complementary DNA
 A;Reference number: JU0238
 A;Accession: JU0238
 A;Molecule type: protein
 A;Residues: 1-134 <NA2>
 C;Keywords: brain

Query Match 49.4%; Score 44; DB 2; Length 134;
 Best Local Similarity 46.2%; Pred. No. 4.2;
 Matches 6; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

Qy 5 QPLQVMMPEQGD 17
 Db 105 EPLIEPLMEPRGE 117

RESULT 8
 156498
 Phosphonuroprotein 14 - rat
 C;Species: Rattus sp. (rat)
 C;Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 05-Nov-1999
 C;Accession: 156498
 R;Tobe, T.; Nakajo, S.; Tanaka, A.; Mito, Y.; Omata, K.; Nakaya, M.; Nakam
 J. Neurochem. 59, 1624-1629, 1992
 A;Title: Cloning and characterization of the cDNA encoding a novel brain-specific 14-kDa
 A;Reference number: 156498; MUID:93019240; PMID:11402909
 A;Accession: 156498
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: mRNA
 A;Residues: 1-137 <RES>
 A;Cross-references: GB:D17764; NID:G409780; PIDN:BAA04610.1; PID:G459235

Query Match 49.4%; Score 44; DB 2; Length 137;
 Best Local Similarity 46.2%; Pred. No. 4.3;
 Matches 6; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

Qy 5 QPLQVMMPEQGD 17
 Db 105 EPLIEPLMEPRGE 117

RESULT 9
 C96534
 Probable Poly-A Binding Protein [imported] - Arabidopsis thaliana
 C;Species: Arabidopsis thaliana (mouse-ear cress)
 C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
 C;Accession: C96534
 R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federpiet, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
 ansen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziali,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A;Reference number: A86141; MUID:21016719; PMID:11130712
 A;Accession: C96534
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-671 <STO>
 A;Cross-references: UNIPROT:Q9FXA2; GB:AE005173; NID:g10120431; PIDN:AAG13056.1; GSPDB:G
 C;Genetics:
 A;Gene: FLJ422.3
 A;Map position: 1
 C;Superfamily: polyadenylate-binding protein; ribonucleoprotein repeat homology

Query Match 49.4%; Score 44; DB 2; Length 671;
 Best Local Similarity 46.7%; Pred. No. 26;

Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 2 QTPQPLQVMMPEQG 16
 Db 514 QQSPPMWQQMHPRG 528

RESULT 10
 AB0784
 conserved hypothetical protein STY2446 [imported] - Salmonella enterica subsp. enterica
 C;Species: Salmonella enterica subsp. enterica serovar Typhi
 A;Note: this species has also been called Salmonella typhi
 C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
 C;Accession: AB0784
 R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
 th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
 S.; Moule, S.; O'Gaora, P.
 Nature 413, 848-852, 2001
 A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
 A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serovar
 A;Reference number: AB0502; MUID:21534947; PMID:11677608
 A;Accession: AB0784
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-134 <PAR>
 A;Cross-references: GB:AL513382; PIDN:CAD02592.1; PID:g16503449; GSPDB:GN00176
 C;Genetics:
 A;Gene: STY2446

Query Match 48.3%; Score 43; DB 2; Length 134;
 Best Local Similarity 61.5%; Pred. No. 6.2;
 Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 2 QTPQPLQVMMPE 14
 Db 63 QAPQPLRQSM LSP 75

RESULT 11
 A60633
 tetracycline resistance protein - Staphylococcus aureus (strain MRS101)
 C;Species: Staphylococcus aureus
 C;Date: 28-Apr-1993 #sequence_revision 28-Apr-1993 #text_change 09-Jul-2004
 C;Accession: A60633
 R;Nesin, M.; Svec, P.; Lupski, J.R.; Godson, G.N.; Kreiswirth, B.; Kornblum, J.; Projan,
 Antimicrob. Agents Chemother. 34, 2273-2276, 1990
 A;Title: Cloning and nucleotide sequence of a chromosomally encoded tetracycline resistance
 A;Reference number: A60633; MUID:91158314; PMID:2073121
 A;Accession: A60633
 A;Molecule type: DNA
 A;Residues: 1-639 <NES>
 A;Cross-references: UNIPROT:Q53770; GB:M21136; NID:g153114; PIDN:AAA26678.1; PID:g153115
 C;Superfamily: translation elongation factor G; translation elongation factor Tu homolog
 C;Keywords: antibiotic resistance; GTP binding; nucleotide binding; P-loop
 F;4-131/Domain: translation elongation factor Tu homology <ETU>
 F;10-17/Region: nucleotide-binding motif A (P-loop)
 F;128-131/Region: GTP-binding NKXD motif
 F;220-222/Region: GTP-binding SAK/L motif
 F;16,17,55,128,129,131,220/Binding site: Mg-GTP (Lys, Thr, Thr, Asn, Lys, Asp, Ser) site

Query Match 48.3%; Score 43; DB 2; Length 639;
 Best Local Similarity 50.0%; Pred. No. 36;
 Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 LQTPQPLQVMMPE 14
 Db 339 IENPHLLQTTVEP 352

RESULT 12
 A56779
 tetracycline resistance protein TetM - Enterococcus faecalis transposon Tn916
 C;Species: Enterococcus faecalis

C>Date: 08-Sep-1995 #sequence_revision 08-Sep-1995 #text_change 09-Jul-2004
C/Accession: A56779
R:Su, Y.A.; He, P.; Clewell, D.B.
Antimicrob. Agents Chemother. 36, 769-778, 1992
A/Title: Characterization of the tet(M) determinant of Tn916: evidence for regulation by
A/Reference number: A56779; MUID:92368175; PMID:1323953
A/Accession: A56779
A/Status: preliminary
A/Module type: DNA
A/Residues: 1-639 <SUI>
A/Cross-references: UNIPROT:Q57224; GB:M85225; NID:g148321; PIDN:AAA24784.1; PID:g148322
A/Note: sequence extracted from NCBI backbone (NCBIN:111076, NCBIP:111078)
C/Superfamily: translation elongation factor G; translation elongation factor Tu homolog
C/Keywords: antibiotic resistance; GTP binding; nucleotide binding; P-loop
F/4-131/Domain: translation elongation factor Tu homology <ETU>
F/10-17/Region: nucleotide-binding motif A (P-loop)
F/128-131/Region: GTP-binding NKXD motif
F/220-222/Region: GTP-binding SAK/L motif
F/16,17,55,128,129,131,220/Binding site: Mg-GTP (Lys, Thr, Asn, Lys, Asp, Ser) #sta

Query Match 48.3%; Score 43; DB 2; Length 639;
Best Local Similarity 50.0%; Pred. No. 36;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 LQTPQLLQVMNEP 14
: : | | | | : :
Db 339 IENPHPLLQTTVEP 352

RESULT 13
S13142
tetracycline resistance protein tetM - Enterococcus faecalis Tn916
C/Species: Enterococcus faecalis
C/Date: 13-Jan-1995 #sequence_revision 10-Feb-1995 #text_change 09-Jul-2004
C/Accession: S13142
R:Burdett, V.
Nucleic Acids Res. 18, 6137, 1990
A/Title: Nucleotide sequence of the tet(M) gene of Tn916.
A/Reference number: S13142; MUID:91045089; PMID:2172929
A/Accession: S13142
A/Module type: DNA
A/Residues: 1-639 <EUR>
A/Cross-references: UNIPROT:P21598; EMBL:X56353; NID:g47061; PIDN:CAA39796.1; PID:g47062
J. Biol. Chem. 266, 2872-2877, 1991
A/Title: Purification and characterization of Tet(M), a protein that renders ribosomes
A/Reference number: A23749; MUID:91131580; PMID:1993661
A/Contents: annotation; function; amino end of purified recombinant protein
A/Note: protein purified after expression in E. coli shown to be ribosome-dependent GTPa
C/Genetics:
A/Gene: tetM
C/Superfamily: translation elongation factor G; translation elongation factor Tu homolog
C/Keywords: antibiotic resistance; GTP binding; nucleotide binding; P-loop
F/4-131/Domain: translation elongation factor Tu homology <ETU>
F/10-17/Region: nucleotide-binding motif A (P-loop)
F/128-131/Region: GTP-binding NKXD motif
F/220-222/Region: GTP-binding SAK/L motif
F/16,17,55,128,129,131,220/Binding site: Mg-GTP (Lys, Thr, Asn, Lys, Asp, Ser) #sta

Query Match 48.3%; Score 43; DB 2; Length 639;
Best Local Similarity 50.0%; Pred. No. 36;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 LQTPQLLQVMNEP 14
: : | | | | : :
Db 339 IENPHPLLQTTVEP 352

RESULT 14
AE1406
translation elongation factor G homolog fus [imported] - Listeria monocytogenes (strain
C/Species: Listeria monocytogenes
C/Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 12, 2005, 07:49:31 ; Search time 80.6727 Seconds
(without alignments)
107.909 Million cell updates/sec

Title: US-10-691-330-2

Perfect score: 89

Sequence: 1 LQTPQLQVMPEQGD 17

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_03.*

1: uniprot_sprot.*

2: uniprot_trenbl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48	53.9	691	1 EFG_STRPY	P82477 streptococc
2	48	53.9	692	1 EFG_STR3	O83e7 streptococc
3	48	53.9	692	1 EFG_STRAS	O8dx8 streptococc
4	48	53.9	693	1 EFG_STRMU	Q8dvw4 streptococc
5	48	53.9	698	2 Q74L90	Q4190 lactobacill
6	48	53.9	709	1 EFG_LACLA	Q9cdg1 lactococcus
7	47.5	53.4	192	2 Q928G4	Q928g4 listeria in
8	46	51.7	473	2 Q84GV7	Q84gv7 lactobacill
9	46	51.7	473	2 Q84GV8	Q84gv8 lactobacill
10	46	51.7	474	2 Q84GV1	Q84gv1 lactobacill
11	46	51.7	474	2 Q84GW2	Q84gw2 lactobacill
12	46	51.7	475	2 Q84GU9	Q84gu9 lactobacill
13	46	51.7	475	2 Q84GW0	Q84gw0 lactobacill
14	46	51.7	477	2 Q84GW3	Q84gw3 lactobacill
15	46	51.7	477	2 Q84GW4	Q84gw4 lactobacill
16	46	51.7	484	2 Q84GU5	Q84gu5 lactobacill
17	46	51.7	639	1 TETM_UREUR	P09757 ureaplasma
18	46	51.7	693	1 EFG_STRPN	P64022 streptococc
19	46	51.7	693	1 EFG_STRR6	P64023 streptococc
20	45	50.6	667	2 Q35745	Q35745 mus musculu
21	45	50.6	675	2 Q35743	Q35743 mus musculu
22	45	50.6	704	2 Q9JLQ4	Q9Jlq4 mus musculu
23	44.5	50.0	198	2 Q7MDC5	Q7mdc5 vibrio vuln
24	44	49.4	133	2 Q1LZ23	Q1lzz3 mus musculu
25	44	49.4	134	1 SYUB_BOVIN	P33567 bos taurus
26	44	49.4	134	1 SYUB_HUMAN	Q16143 homo sapien
27	44	49.4	137	1 SYUB_RAT	Q63754 rattus norv
28	44	49.4	172	2 Q6CMX4	Q6cmx4 kluyveromyc
29	44	49.4	182	2 Q9U558	Q9u558 plasmodium
30	44	49.4	336	2 Q8W424	Q8w424 gossypium h
31	44	49.4	348	2 Q7Q8X8	Q7q8x8 anopheles g

32	44	49.4	366	2 Q89Y61	Q89y61 bradyrhizob
33	44	49.4	368	2 Q6V9S8	Q6v9s8 gossypium h
34	44	49.4	418	2 Q8TA00	Q8ta00 trypanosoma
35	44	49.4	622	2 Q8TA01	Q8ta01 trypanosoma
36	44	49.4	671	2 Q9FXA2	Q9fxa2 arabidopsis
37	44	49.4	829	2 Q7YW24	Q7yw24 trypanosoma
38	44	49.4	833	2 Q8IF10	Q8if10 trypanosoma
39	44	49.4	833	2 Q8IF12	Q8if12 trypanosoma
40	44	49.4	834	2 Q8T9M8	Q8t9m8 trypanosoma
41	44	49.4	1058	2 Q7T5C7	Q7t5c7 cercopithec
42	44	49.4	1058	2 Q805X6	Q805x6 cercopithec
43	44	49.4	1232	2 Q7SAH0	Q7sah0 neurospora
44	44	49.4	1266	2 Q633A0	Q633a0 bacillus ce
45	44	49.4	2181	2 Q7M561	Q7m561 brachydanio

ALIGNMENTS

RESULT 1

ID	EFG_STRPY	STANDARD	PRT	691 AA
AC	P82477			
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	25-OCT-2004 (Rel. 45, Last annotation update)			
DE	Elongation factor G (EF-G).			
GN	Name=fus; Synonyms=fusA;			
GN	OrderedLocusNames=Spy0273, SpyM3_0200, SP80206, spyM18_0260;			
OS	Streptococcus pyogenes,			
OS	Streptococcus pyogenes (serotype M3), and			
OS	Streptococcus pyogenes (serotype M18).			
OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;			
OC	Streptococcus.			
OX	NCBI_taxid=1314, 198466, 186103;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=SF370 / ATCC 700294 / Serotype M1;			
RX	MEDLINE=21192684; PubMed=11296296; DOI=10.1073/pnas.071559398;			
RA	Ferretti J.J., McShan W.M., Ajdic D.J., Savic G., Lyon K.,			
RA	Primeaux C., Setate S., Suvoarov A.N., Kenton S., Lai H.S., Lin S.P.,			
RA	Qian Y., Jia H.G., Najaf F.Z., Ren Q., Zhu H., Song L., White J.,			
RA	Yuan X., Clifton S.W., Roe B.A., McLaughlin R.E.;			
RA	"Complete genome sequence of an M1 strain of Streptococcus pyogenes.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=MGAS315 / Serotype M3;			
RX	MEDLINE=22133808; PubMed=12122206; DOI=10.1073/pnas.152298499;			
RA	Beres S.B., Sylva G.L., Barbican K.D., Lei B., Hoff J.S.,			
RA	Mamarella N.D., Liu M.-Y., Smoot J.C., Porcella S.F., Perkins L.D.,			
RA	Campbell D.S., Smith T.M., McCormick J.K., Leung D.Y.M.,			
RA	Schlievert P.M., Mueser J.M.;			
RA	"Genome sequence of a serotype M3 strain of group A Streptococcus:			
RT	phage-encoded toxins, the high-virulence phenotype, and clone			
RT	emergence.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 99:10078-10083(2002).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=SSI-1 / Serotype M3;			
RX	MEDLINE=22683278; PubMed=12799345; DOI=10.1101/gr.1096703;			
RA	Nakagawa I., Kurokawa K., Yamashita A., Nakata M., Tomiyasu Y.,			
RA	Okahashi N., Kawabata S., Yamazaki K., Shiba T., Yasunaga T.,			
RA	Hayashi H., Hattori M., Hamada S.;			
RT	"Genome sequence of an M3 strain of Streptococcus pyogenes reveals a			
RT	large-scale genomic rearrangement in invasive strains and new insights			
RT	into phage evolution.";			
RL	Genome Res. 13:1042-1055(2003).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=MGAS232 / Serotype M18;			
RX	PubMed=11917108; DOI=10.1073/pnas.062526099;			
RA	Smoot J.C., Barbican K.D., Van Gompel J.J., Smoot L.M., Chaussee M.S.,			

RA Sylva G.L., Sturdevant D.E., Ricklefs S.M., Porcella S.F.,
 RA Parkins L.D., Beres S.B., Campbell D.S., Smith T.M., Zhang Q.,
 RA Kapur V., Daly J.A., Veasy L.G., Musser J.M.;
 RT "Genome sequence and comparative microarray analysis of serotype M18
 RT group A Streptococcus strains associated with acute rheumatic fever
 RT outbreaks";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:4668-4673(2002).
 RN (5)
 RP PARTIAL SEQUENCE, AND MASS SPECTROMETRY.
 RC STRAIN=JRS4 / Serotype M6;
 RA Hogan D.A., Du P., Stevenson T.I., Whitton M., Kilby G.W., Rogers J.,
 RA VanBogelen R.A.;
 RT "Two-dimensional gel electrophoresis map of Streptococcus pyogenes
 RT proteins";
 RL Submitted (MAY-2000) to Swiss-Prot.
 CC -1- FUNCTION: This protein promotes the GTP-dependent translocation of
 CC the nascent protein chain from the A-site to the P-site of the
 CC ribosome (By similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: Belongs to the GTP-binding elongation factor family.
 CC EF-G/EF-2 subfamily.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; A006493; AAK33347.1; -
 DR EMBL; AE014140; AAM78807.1; -
 DR EMBL; AP005141; BAC63301.1; -
 DR EMBL; AE009973; AAL97040.1; -
 DR HSSP; P13551.1DAR.
 DR HAMAP; MF_00054; -; 1.
 DR InterPro; IPR004540; EF-G.
 DR InterPro; IPR000640; EFG_C.
 DR InterPro; IPR009022; EFG_III_V.
 DR InterPro; IPR005517; EFG_IV.
 DR InterPro; IPR004161; EFTU_D2.
 DR InterPro; IPR000795; ProtSyn_GTPbind.
 DR InterPro; IPR005225; Small_GTP.
 DR Pfam; PF00679; EFG_C; 1.
 DR Pfam; PF03764; EFG_IV; 1.
 DR Pfam; PF00009; GTP_EFTU; 1.
 DR Pfam; PF03144; GTP_EFTU_D2; 1.
 DR PRINTS; PR00315; ELONGATNFCT.
 DR TIGRFAMS; TIGR00484; EF-G; 1.
 DR PROSITE; PS00301; EFACOR_GTP; 1.
 DR Complete proteome; Direct protein sequencing; Elongation factor;
 DR GTP-binding; Protein biosynthesis.
 FT INIT_MET 0
 FT NP_BIND 16 23 GTP (By similarity).
 FT NP_BIND 80 84 GTP (By similarity).
 FT NP_BIND 134 137 GTP (By similarity).
 SQ SEQUENCE 691 AA; 76397 MW; D59B857A2CDB40CD CRC64;
 Query Match 53.9%; Score 48; DB 1; Length 691;
 Best Local Similarity 40.0%; Pred. No. 35;
 Matches 6; Conservative 8; Mismatches 1; Indels 0; Gaps 0;
 QY 1 LOTPQPLLQVMVEPQ 15
 Db 399 IEVPEPVIQLMVEPK 413
 RESULT 2
 EFG_STRAS3 STANDARD; PRT; 692 AA.
 AC QBE3E7;

DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Elongation factor G (EF-G).
 GN Name=fusa; OrderedLocusNames=gbs1812;
 OS Streptococcus agalactiae (serotype III).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=216495;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MEM316 / Serotype III;
 RX MEDLINE=22242508; PubMed=12354221;
 RA Glaser P., Rusniok C., Buchrieser C., Chevalier P., Frangeul L.,
 RA Maadek T., Zouine M., Couve E., Lalioui L., Poyart C., Trieu-Cuot P.,
 RA Kunst F.;
 RT "Genome sequence of Streptococcus agalactiae, a pathogen causing
 RT invasive neonatal disease";
 RL Mol. Microbiol. 45:1499-1513(2002).
 CC -1- FUNCTION: This protein promotes the GTP-dependent translocation of
 CC the nascent protein chain from the A-site to the P-site of the
 CC ribosome.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: Belongs to the GTP-binding elongation factor family.
 CC EF-G/EF-2 subfamily.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; AL766853; CAD47471.1; -
 DR HSSP; P13551; 1DAR.
 DR Sagalists; gbs1812; -
 DR HAMAP; MF_00054; -; 1.
 DR InterPro; IPR004540; EF-G.
 DR InterPro; IPR000640; EFG_C.
 DR InterPro; IPR009022; EFG_III_V.
 DR InterPro; IPR005517; EFG_IV.
 DR InterPro; IPR004161; EFTU_D2.
 DR InterPro; IPR000795; ProtSyn_GTPbind.
 DR InterPro; IPR005225; Small_GTP.
 DR Pfam; PF00679; EFG_C; 1.
 DR Pfam; PF03764; EFG_IV; 1.
 DR Pfam; PF00009; GTP_EFTU; 1.
 DR Pfam; PF03144; GTP_EFTU_D2; 1.
 DR PRINTS; PR00315; ELONGATNFCT.
 DR TIGRFAMS; TIGR00484; EF-G; 1.
 DR PROSITE; PS00301; EFACOR_GTP; 1.
 DR Complete proteome; Elongation factor; GTP-binding;
 DR Protein biosynthesis.
 FT NP_BIND 17 24 GTP (By similarity).
 FT NP_BIND 81 85 GTP (By similarity).
 FT NP_BIND 135 138 GTP (By similarity).
 SQ SEQUENCE 692 AA; 76598 MW; 199DA941BC7F9182 CRC64;
 Query Match 53.9%; Score 48; DB 1; Length 692;
 Best Local Similarity 40.0%; Pred. No. 35;
 Matches 6; Conservative 8; Mismatches 1; Indels 0; Gaps 0;
 QY 1 LOTPQPLLQVMVEPQ 15
 Db 400 IEVPEPVIQLMVEPK 414
 RESULT 3
 EFG_STRAS5 STANDARD; PRT; 692 AA.
 ID_EFG_STRAS5

Q8DXS7;
 10-OCT-2003 (Rel. 42, Created)
 10-OCT-2003 (Rel. 42, Last sequence update)
 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Elongation factor G (EF-G).
 GN Name=fusA; OrderedLocuNames=SAG1769;
 OS Streptococcus agalactiae (serotype V).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=216466;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=2603 V/R / Serotype V;
 RX MEDLINE=2222988; PubMed=12200547; DOI=10.1073/pnas.182380799;
 RA Tettelin H., Maignani V., Cieslewicz M.J., Eisen J.A., Peterson S.N.,
 RA Wessels M.R., Paulsen I.T., Nelson K.E., Margarit I., Read T.D.,
 RA Madoff L.C., Wolf A.M., Beanan M.J., Brinkac L.M., Daugherty S.C.,
 RA DeBoy R.T., Wolf A.S., Kolonay J.F., Madupu R., Lewis M.R.,
 RA Radu D., Fedorova N.B., Scanlan D., Khouri H.M., Mulligan S.,
 RA Carty H.A., Cline R.T., Van Aken S.E., Gill J., Scarselli M., Mora M.,
 RA Iacobini E.T., Brettoni C., Galli G., Mariani M., Vegni F., Malone D.,
 RA Rinaudo D., Rappuoli R., Telford J.L., Kasper D.L., Grandi G.,
 RA Fraser C.M.;
 RT "Complete genome sequence and comparative genomic analysis of an
 emerging human pathogen, serotype V Streptococcus agalactiae.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:12391-12396(2002).
 CC -1- FUNCTION: This protein promotes the GTP-dependent translocation of
 the nascent protein chain from the A-site to the P-site of the
 ribosome.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: Belongs to the GTP-binding elongation factor family.
 CC EF-G/EF-2 subfamily.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AE014272; AAN00632.1; -
 CC HSSP; P13551; 1DAR.
 CC TIGR; SAG1769; -; 1
 CC HAMAP; MF_00054; -; 1
 CC InterPro; IPR004540; EF-G.
 CC InterPro; IPR000640; EF-G.
 CC InterPro; IPR009022; EFG_III_V.
 CC InterPro; IPR005517; EFG_IV.
 CC InterPro; IPR004161; EFTU_D2.
 CC InterPro; IPR000795; ProtSyn_GTPbind.
 CC InterPro; IPR005225; Small_GTP.
 CC Pfam; PF00679; EFG_C; 1.
 CC Pfam; PF03764; EFG_IV; 1.
 CC Pfam; PF00009; GTP_EFTU; 1.
 CC Pfam; PF03144; GTP_EFTU_D2; 1.
 CC PRINTS; PR00315; ELONGATNFCT.
 CC TIGRFAMs; TIGR00484; EF-G; 1.
 CC TIGRFAMs; TIGR00231; small_GTP; 1.
 CC PROSITE; PS00301; EFACOR_GTP; 1.
 CC Complete proteome; Elongation factor; GTP-binding;
 KW Protein biosynthesis.
 KW NP_BIND 17 24 GTP (By similarity).
 FT NP_BIND 81 85 GTP (By similarity).
 FT NP_BIND 135 138 GTP (By similarity).
 FT NP_BIND 135 138 GTP (By similarity).
 SQ SEQUENCE 692 AA; 76570 MW; 36BCF199BA006AB1 CRC64;

Query Match 53.9%; Score 48; DB 1; Length 692;
 Best Local Similarity 40.0%; Pred. No. 35;
 Matches 6; Conservative 8; Mismatches 1; Indels 0; Gaps 0;
 1 LQTPQLQVMEPK 15

Db 400 IEVPEPVLQVMEPK 414
 RESULT 4
 EFG_STRMU STANDARD; PRT; 693 AA.
 ID Q8DV4; 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Elongation factor G (EF-G).
 GN Name=fusA; OrderedLocuNames=SMU_359;
 OS Streptococcus mutans.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1309;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=UA159 / ATCC 700610 / Serotype C;
 RX MEDLINE=22295063; PubMed=12397186; DOI=10.1073/pnas.172501299;
 RA Ajdic D.J., McShan W.M., McLaughlin R.E., Savic G., Chang J.,
 RA Carson M.B., Primeaux C., Tian R., Kenton S., Jia H.G., Lin S.P.,
 RA Qian Y., Li S., Zhu H., Majar F.Z., Lai H., White J., Roe B.A.,
 RA Ferretti J.J.;
 RT "Genome sequence of Streptococcus mutans UA159, a cariogenic dental
 pathogen.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).
 CC -1- FUNCTION: This protein promotes the GTP-dependent translocation of
 the nascent protein chain from the A-site to the P-site of the
 ribosome.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: Belongs to the GTP-binding elongation factor family.
 CC EF-G/EF-2 subfamily.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AE014883; AAN58117.1; -
 CC HSSP; P13551; 1DAR.
 CC HAMAP; MF_00054; -; 1
 CC InterPro; IPR004540; EF-G.
 CC InterPro; IPR000640; EF-G.
 CC InterPro; IPR009022; EFG_III_V.
 CC InterPro; IPR005517; EFG_IV.
 CC InterPro; IPR004161; EFTU_D2.
 CC InterPro; IPR000795; ProtSyn_GTPbind.
 CC InterPro; IPR005225; Small_GTP.
 CC Pfam; PF00679; EFG_C; 1.
 CC Pfam; PF03764; EFG_IV; 1.
 CC Pfam; PF00009; GTP_EFTU; 1.
 CC Pfam; PF03144; GTP_EFTU_D2; 1.
 CC PRINTS; PR00315; ELONGATNFCT.
 CC TIGRFAMs; TIGR00484; EF-G; 1.
 CC TIGRFAMs; TIGR00231; small_GTP; 1.
 CC PROSITE; PS00301; EFACOR_GTP; 1.
 CC Complete proteome; Elongation factor; GTP-binding;
 KW Protein biosynthesis.
 KW NP_BIND 17 24 GTP (By similarity).
 FT NP_BIND 81 85 GTP (By similarity).
 FT NP_BIND 135 138 GTP (By similarity).
 FT NP_BIND 135 138 GTP (By similarity).
 SQ SEQUENCE 693 AA; 76664 MW; 252763A3606B5E04 CRC64;

Query Match 53.9%; Score 48; DB 1; Length 693;
 Best Local Similarity 40.0%; Pred. No. 35;
 Matches 6; Conservative 8; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 LQTPQLLQVMMPEQ 15
Db 400 IEVPEPVIQVMVEPK 414

RESULT 5
Q74L90
ID Q74L90 PRELIMINARY; PRT; 698 AA.
AC Q74L90;
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DE Elongation factor G.
GN OrderedLocusNames=LJ0337;
OS Lactobacillus johnsonii.
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC Lactobacillus.
OX NCBI_TaxID=33959;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCC 533;
RX PubMed=14966310;
RA Pridmore R.D., Berger B., Desiere F., Vilanova D., Barretto C.,
RA Pittet A.-C., Zwielen M.-C., Rouvet M., Altermann E., Barrangou R.,
RA Mollet B., Mercenier A., Klaenhammer T., Arigoni F., Schell M.A.;
RT "The genome sequence of the probiotic intestinal bacterium
RT Lactobacillus johnsonii NCC 533."
RL Proc. Natl. Acad. Sci. U.S.A. 101:2512-2517(2004).
DR EMBL; AE017201; AAS08323.1; -.
DR GO; GO:0005252; F:GTP binding; IEA.
DR GO; GO:0003746; P:translation elongation factor activity; IEA.
DR GO; GO:0006412; P:protein biosynthesis; IEA.
DR GO; GO:0006414; P:translational elongation; IEA.
DR InterPro; IPR009063; Bac_Ig/alb_bind.
DR InterPro; IPR004540; EF-G.
DR InterPro; IPR000640; EFG C.
DR InterPro; IPR009022; EFG III V.
DR InterPro; IPR005517; EFG IV.
DR InterPro; IPR004161; EFTU D2.
DR InterPro; IPR000795; ProtSyn GTPbind.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF00679; EFG_C; 1.
DR Pfam; PF03764; EFG_IV; 1.
DR Pfam; PF00009; GTP_EFTU; 1.
DR Pfam; PF03144; GTP_EFTU_D2; 1.
DR PRINTS; PR00315; ELONGATNFCT.
DR TIGRFAMS; TIGR00484; EF-G; 1.
DR TIGRFAMS; TIGR00231; small_GTP; 1.
DR PROSITE; PS00301; EFATOR_GTP; 1.
KW Complete proteome; Elongation factor; GTP-binding.
SQ SEQUENCE 698 AA; 76916 MW; B73C135DE5F86E2 CRC64;

Query Match 53.9%; Score 48; DB 2; Length 698;
Best Local Similarity 53.3%; Pred. No. 36;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 LQTPQLLQVMMPEQ 15
Db 403 LQVDPVIQVSVEPK 417

RESULT 6
EFG_LACLA
ID EFG_LACLA STANDARD; PRT; 709 AA.
AC O9CDG1;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE Elongation factor G (EF-G).
GN Name=fusA; OrderedLocusNames=LL2260;
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.

QY 1 LQTPQLLQVMMPEQ 15
Db 415 IEVPEPVIQVMVEPK 429

RESULT 7
Q928G4
ID Q928G4 PRELIMINARY; PRT; 192 AA.
AC Q928G4;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Lin2571 protein.
GN OrderedLocusNames=lin2571;
OS Listeria innocua.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX NCBI_TaxID=1642;

NCBI_TaxID=1360;
[1]
RN RP SEQUENCE FROM N.A.
RC STRAIN=IL1403;
RX MEDLINE=21235186; PubMed=11337471; DOI=10.1101/gr.169701;
RA Boletín A., Wincker P., Mauger S., Jaillon O., Malarne K.,
RA Weissenbach J., Ehrlich S.D., Sorokin A.;
RT "The complete genome sequence of the lactic acid bacterium Lactococcus
RT lactis ssp. lactis IL1403."
RL Genome Res. 11:731-753(2001).
CC -!- FUNCTION: This protein promotes the GTP-dependent translocation of
CC the nascent protein chain from the A-site to the P-site of the
CC ribosome.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to the GTP-binding elongation factor family.
CC EF-G/EF-2 subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE006455; AAK06358.1; -.
DR PIR; D86907; D86907.
DR HSSP; P13551; IDAR.
DR HAMAP; MF 00054; -.
DR InterPro; IPR004540; EF-G.
DR InterPro; IPR000640; EFG C.
DR InterPro; IPR009022; EFG III V.
DR InterPro; IPR005517; EFG IV.
DR InterPro; IPR004161; EFTU D2.
DR InterPro; IPR000795; ProtSyn GTPbind.
DR InterPro; IPR005225; Small_GTP.
DR InterPro; IPR009000; Translat_factor.
DR Pfam; PF00679; EFG_C; 1.
DR Pfam; PF03764; EFG_IV; 1.
DR Pfam; PF00009; GTP_EFTU; 1.
DR Pfam; PF03144; GTP_EFTU_D2; 1.
DR PRINTS; PR00315; ELONGATNFCT.
DR TIGRFAMS; TIGR00484; EF-G; 1.
DR TIGRFAMS; TIGR00231; small_GTP; 1.
DR PROSITE; PS00301; EFATOR_GTP; 1.
KW Complete proteome; Elongation factor; GTP-binding;
KW Protein biosynthesis.
FT NP_BIND 17 24 GTP (By similarity).
FT NP_BIND 81 85 GTP (By similarity).
FT NP_BIND 135 138 GTP (By similarity).
SQ SEQUENCE 709 AA; 77955 MW; 1EAD29D65A748032 CRC64;

Query Match 53.9%; Score 48; DB 1; Length 709;
Best Local Similarity 40.0%; Pred. No. 36;
Matches 6; Conservative 8; Mismatches 1; Indels 0; Gaps 0;

QY 1 LQTPQLLQVMMPEQ 15
Db 415 IEVPEPVIQVMVEPK 429

RESULT 7
Q928G4
ID Q928G4 PRELIMINARY; PRT; 192 AA.
AC Q928G4;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Lin2571 protein.
GN OrderedLocusNames=lin2571;
OS Listeria innocua.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX NCBI_TaxID=1642;

```

```

RN  SEQUENCE FROM N.A.
PC  STRAIN-CLIP 11262 / Serovar 6a;
RX  MEDLINE=21537279; PubMed=11679669; DOI=10.1126/science.1063447;
RA  Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
RA  Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
RA  Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
RA  Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA  Entian K.-D., Fajhi H., Garcia-del Portillo F., Garrido P.,
RA  Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA  Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapkai G.,
RA  Madueno E., Maikutnam A., Mata Vicente J., Ng E., Nedjari H.,
RA  Nordstiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA  Remmel B., Rose M., Schluter T., Simoes N., Tierrez A.,
RA  Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
RT  "Comparative genomics of Listeria species.";
RL  Science 294:849-852(2001).
DR  EMBL; AF1753; AF1753.
DR  PIR; AF1753; AF1753.
DR  Listalist; LIN2571; -.
DR  TIGRFAMs; TIGR01603; maj_tail_phil3; 1.
KW  Complete proteome.
SQ  SEQUENCE 192 AA; 20841 MW; BB4B4D3261DDA74 CRC64;

Query Match 53.4%; Score 47.5; DB 2; Length 192;
Best Local Similarity 50.0%; Pred. No. 11;
Matches 9; Conservative 5; Mismatches 1; Indels 3; Gaps 1;

Qy 3 TPQPL---LQVMPQGD 17
Db 31 TPKPLPGGVETSVQGD 48

RESULT 8
Q84GV7
ID Q84GV7 PRELIMINARY; PRT; 473 AA.
AC Q84GV7
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Tetracycline resistance protein (Fragment).
GN Name=tetM;
OS Lactobacillus sakei subsp. sakei.
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC Lactobacillus.
OX NCBI_TaxID=214326;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22458139; PubMed=12571056;
RX DOI=10.1128/AEM.69.2.1270-1275.2003;
RA Gevers D., Danielsen M., Huys G., Swings J.;
RT "Molecular characterization of tet(M) genes in Lactobacillus isolates
from different types of fermented dry sausage.";
RL Appl. Environ. Microbiol. 69:1270-1275(2003).
CC -!- FUNCTION: Abolishes the inhibitory effect of tetracycline on
protein synthesis by a non-covalent modification of the ribosomes
(BY similarity).
CC -!- SIMILARITY: Belongs to the GTP-binding elongation factor family.
CC TetM/tetO subfamily.
DR EMBL; AY149583; AAN84487.1; -.
DR GO; GO:0005525; P:GTP binding; IEA.
DR GO; GO:0003746; P:translation elongation factor activity; IEA.
DR GO; GO:0006412; P:protein biosynthesis; IEA.
DR GO; GO:0046677; P:response to antibiotic; IEA.
DR GO; GO:0006414; P:translational elongation; IEA.
DR InterPro; IPR000640; EFG_C.
DR InterPro; IPR009022; EFG_III_V.
DR InterPro; IPR005517; EFG_IV.
DR InterPro; IPR004161; EFTU_D2.
DR InterPro; IPR000795; ProtSyn_GTPbind.
DR InterPro; IPR002127; TCR_TetOQM.
DR InterPro; IPR009000; Translat_factor.
DR Pfam; PF00679; EFG_C; 1.
DR Pfam; PF03764; EFG_IV; 1.
DR Pfam; PF00009; GTP_EFTU; 1.
DR Pfam; PF03144; GTP_EFTU_D2; 1.
DR PRINTS; PR01037; TCR_TetOQM.
KW Antibiotic resistance; Cyclin; GTP-binding; Protein biosynthesis.
FT NON_TER 473
FT NON_TER 473
SQ SEQUENCE 473 AA; 53962 MW; 3F7AA0C3F577C133 CRC64;

Query Match 51.7%; Score 46; DB 2; Length 473;
Best Local Similarity 50.0%; Pred. No. 50;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 1 LQTPQLLQVMP 14
Db 246 IENPHPLQLITV 259

RESULT 9
Q84GV8
ID Q84GV8 PRELIMINARY; PRT; 473 AA.
AC Q84GV8
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Tetracycline resistance protein (Fragment).
GN Name=tetM;
OS Lactobacillus sakei subsp. carnosus.
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC Lactobacillus.
OX NCBI_TaxID=214325;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22458139; PubMed=12571056;
RX DOI=10.1128/AEM.69.2.1270-1275.2003;
RA Gevers D., Danielsen M., Huys G., Swings J.;
RT "Molecular characterization of tet(M) genes in Lactobacillus isolates
from different types of fermented dry sausage.";
RL Appl. Environ. Microbiol. 69:1270-1275(2003).
CC -!- FUNCTION: Abolishes the inhibitory effect of tetracycline on
protein synthesis by a non-covalent modification of the ribosomes
(BY similarity).
CC -!- SIMILARITY: Belongs to the GTP-binding elongation factor family.
CC TetM/tetO subfamily.
DR EMBL; AY149583; AAN84487.1; -.
DR GO; GO:0005525; P:GTP binding; IEA.
DR GO; GO:0003746; P:translation elongation factor activity; IEA.
DR GO; GO:0006412; P:protein biosynthesis; IEA.
DR GO; GO:0046677; P:response to antibiotic; IEA.
DR GO; GO:0006414; P:translational elongation; IEA.
DR InterPro; IPR000640; EFG_C.
DR InterPro; IPR009022; EFG_III_V.
DR InterPro; IPR005517; EFG_IV.
DR InterPro; IPR004161; EFTU_D2.
DR InterPro; IPR000795; ProtSyn_GTPbind.
DR InterPro; IPR002127; TCR_TetOQM.
DR InterPro; IPR009000; Translat_factor.
DR Pfam; PF00679; EFG_C; 1.
DR Pfam; PF03764; EFG_IV; 1.
DR Pfam; PF00009; GTP_EFTU; 1.
DR Pfam; PF03144; GTP_EFTU_D2; 1.
DR PRINTS; PR01037; TCR_TetOQM.
KW Antibiotic resistance; Cyclin; GTP-binding; Protein biosynthesis.
FT NON_TER 473
FT NON_TER 473
SQ SEQUENCE 473 AA; 53962 MW; 3F7AA0C3F577C133 CRC64;

Query Match 51.7%; Score 46; DB 2; Length 473;
Best Local Similarity 50.0%; Pred. No. 50;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 1 LQTPQLLQVMP 14
Db 246 IENPHPLQLITV 259
```


10

Search completed: April 12, 2005, 08:05:37
Job time : 82.6727 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 12, 2005, 07:49:31 ; Search time 100.8 Seconds
(without alignments)
69.064 Million cell updates/sec

Title: US-10-691-330-3

Perfect score: 98

Sequence: 1 DQPPDVEKPLQPFQVQS 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq_16Dec04:*

- 1: Geneseqp1980s:*
- 2: Geneseqp1990s:*
- 3: Geneseqp2000s:*
- 4: Geneseqp2001s:*
- 5: Geneseqp2002s:*
- 6: Geneseqp2003as:*
- 7: Geneseqp2003bs:*
- 8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	98	100.0	18	4	Aab72502 Colostrin
2	98	100.0	18	4	Aab59325 Ewe colos
3	98	100.0	18	4	Aab72248 Colostrin
4	98	100.0	18	4	Aab72534 Colostrin
5	98	100.0	18	5	Aao14579 Neural ce
6	98	100.0	18	5	Aam51038 Colostrin
7	98	100.0	18	5	Aae20230 Colostrin
8	98	100.0	18	8	Adn60297 Constitue
9	98	100.0	18	8	Ads74401 Ovine col
10	98	100.0	19	4	Aab59355 Ewe colos
11	51	52.0	1047	6	Abu28927 Protein e
12	51	52.0	1056	7	Adh87201 Enterococ
13	48	49.0	116	8	Adr08456 Human pro
14	48	49.0	329	8	Adj71952 Human PMW
15	48	49.0	377	4	Aam79068 Human pro
16	48	49.0	377	5	Abb83472 Human cyt
17	48	49.0	377	8	Adc47756 Human NOV
18	48	49.0	377	8	Adj79026 Human NOV
19	48	49.0	400	7	Adh86504 Enterococ
20	48	49.0	541	4	Aau35153 Enterococ
21	47	48.0	516	5	Abb91350 Herbicida
22	47	48.0	719	5	Abb91351 Herbicida
23	46.5	47.4	120	3	Aag56012 Arabidops
24	46.5	47.4	126	3	Aag56011 Arabidops
25	46.5	47.4	142	3	Aag56010 Arabidops

26 46 46.9 176 4 AAU55015 Propionib
27 46 46.9 176 6 ABM51534 Propionib
28 46 46.9 176 6 ABM64933 Propionib
29 46 46.9 180 3 AAY45073 Rat amelo
30 46 46.9 194 3 AAY45072 Rat amelo
31 45 45.9 26 3 AAB20811 PEST sequ
32 45 45.9 181 4 ABG29328 Novel hum
33 45 45.9 188 4 AAB63292 Human bre
34 45 45.9 188 4 AAB63288 Human bre
35 45 45.9 215 4 AAB63316 Human bre
36 45 45.9 233 4 AAB63317 Human bre
37 45 45.9 325 8 ADP04221 Human col
38 45 45.9 383 4 ABG27213 Novel hum
39 45 45.9 383 4 ABG09426 Human hos
40 45 45.9 620 8 ADH09924 Human hos
41 45 45.9 814 8 ADH09922 Human hos
42 45 45.9 814 8 ADH09923 Human hos
43 45 45.9 814 8 ADL83154 Human PRO
44 45 45.9 814 8 ADP54647 Human PRO
45 45 45.9 814 8 ADP24002 PRO polyp

ALIGNMENTS

RESULT 1
AAB72502
ID AAB72502 standard; peptide; 18 AA.
XX AC AAB72502;
XX AC
DT 09-MAY-2001 (first entry)
XX AC
DE Colostrin peptide #3.
XX AC
KW Dermatological; oxidative stress regulator; colostrin.
XX OS
XX Unidentified.
XX PN W0200112650-A2.
XX PD 22-FEB-2001.
XX PF 17-AUG-2000; 2000WO-US022665.
XX PR 17-AUG-1999; 99US-0149310P.
XX PA (TEXA) UNIV TEXAS SYSTEM.
XX Stanton GU, Hughes TK, Boldogh I;
XX WPI; 2001-218342/22.
XX Modulating oxidative stress level in a cell, involves contacting the cell with an oxidative stress regulator selected from colostrin, its constituent peptide, analog or their combinations.
XX Claim 6; Page 25; 48pp; English.
XX The present invention relates to a method for modulating the oxidative stress level in a cell or a patient, comprising contacting the cell with, or administering to the patient, an oxidative stress regulator selected from colostrin, or its constituent peptide (e.g. the present peptide), to change the level of an oxidising species in the cell. The method can be used to treat oxidative damage to skin, by decreasing or preventing an increase in the level of damage to a biomolecule of the patient
SQ Sequence 18 AA;

Query Match 100.0%; Score 98; DB 4; Length 18;
Best Local Similarity 100.0%; Pred. No. 3.7e-07;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DQPPDVEKPDLPQFQVQS 18
 DB 1 DQPPDVEKPDLPQFQVQS 18

RESULT 2

AAB59325
 ID AAB59325 standard; peptide; 18 AA.

XX AC AAB59325;

XX DT 21-MAR-2001 (first entry)

XX DE Ewe colostrinin peptide fragment B-10.

XX KW Sheep; colostrinin; proline rich polypeptide; colostrum; immune disorder;
 KW central nervous system disorder; dietary supplement; beta-amyloid plaque.

XX OS Ovis sp.

XX PN WO200075173-A2.

XX PD 14-DEC-2000.

XX PF 02-JUN-2000; 2000WO-GB002128.

XX PR 02-JUN-1999; 99GB-00012852.

XX PA (REG-) REGEN THERAPEUTICS PLC.

XX PI Georgiades JA;

XX DR WPI; 2001-071058/08.

XX PT Peptides having an N-terminal amino acid sequence isolated from

PT colostrinin for treating e.g. disorders of the central nervous system and
 PT immune system, viral and bacterial infections, and diseases characterized
 PT by amyloid plaques.

PS Claim 7; Page 27; 63pp; English.

CC The present invention provides the sequences of a number of peptides
 CC found in ewe's colostrinin. Colostrinin is the proline-rich polypeptide
 CC fragment of colostrum. These peptides can be used in the treatment of
 CC central nervous system disorders such as senile dementia, Parkinson's
 CC disease, Alzheimer's disease, psychosis and neurosis, immune system
 CC disorders such as bacterial and viral infections, to improve the
 CC development of a child's immune system, as a dietary supplement, and to
 CC promote the dissolution of beta-amyloid plaques

XX SQ Sequence 18 AA;

Query Match 100.0%; Score 98; DB 4; Length 18;
 Best Local Similarity 100.0%; Pred. No. 3.7e-07;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DQPPDVEKPDLPQFQVQS 18
 DB 1 DQPPDVEKPDLPQFQVQS 18

RESULT 3

AAB72248

ID AAB72248 standard; peptide; 18 AA.

XX AC AAB72248;

XX DT 14-MAY-2001 (first entry)

XX DE Colostrinin derived cytokine inducing peptide SEQ ID 3.

XX KW Colostrinin; immune response; cytokine; blood cell proliferation;
 KW central nervous system disorder; neurological disorder; mental disorder;

XX KW dementia; neurodegenerative disease; Alzheimer's disease; psychosis;
 KW neurosis; infection.

XX OS Synthetic.

XX PN WO200111937-A2.

XX PD 22-FEB-2001.

XX PF 17-AUG-2000; 2000WO-US022818.

XX PR 17-AUG-1999; 99US-0149311P.

XX PA (TEXA) UNIV TEXAS SYSTEM.

XX PA (REG-) REGEN THERAPEUTICS PLC.

XX PI Stanton GJ, Hughes TK, Boldogh I, Georgiades J;

XX DR WPI; 2001-202804/20.

XX PT Inducing a cytokine and modulating an immune response, useful for
 PT treating central nervous system diseases and bacterial and viral
 PT infections, comprises administering colostrinin as an immunological
 PT regulator.

PS Claim 1; Page 34; 50pp; English.

XX CC Sequences AAB72246 - AAB72275 represent peptides derived from colostrinin,
 CC a proline rich polypeptide aggregate contained in colostrum. The peptides
 CC have immune response modulatory activity, and are capable of inducing
 CC cytokines. Colostrinin and its derived peptides are useful for inducing
 CC cytokine production, for modulating an immunological response and for
 CC inducing blood cell proliferation. The peptides are useful in the
 CC treatment of disorders of the central nervous system, neurological
 CC disorders, mental disorders, dementia, neurodegenerative diseases,
 CC Alzheimer's disease, motor neurone disease, psychosis, neurosis, chronic
 CC disorders of the immune system, bacterial and viral infections and
 CC acquired immunological deficiencies

XX SQ Sequence 18 AA;

Query Match 100.0%; Score 98; DB 4; Length 18;
 Best Local Similarity 100.0%; Pred. No. 3.7e-07;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DQPPDVEKPDLPQFQVQS 18
 DB 1 DQPPDVEKPDLPQFQVQS 18

RESULT 4

AAB72534

ID AAB72534 standard; peptide; 18 AA.

XX AC AAB72534;

XX DT 09-MAY-2001 (first entry)

XX DE Colostrinin peptide #3.

XX KW Neuroprotective; neural cell differentiation regulator; colostrinin;
 KW colostrum.

XX OS Unidentified.

XX PN WO200112651-A2.

XX PD 22-FEB-2001.

XX PF 17-AUG-2000; 2000WO-US022774.

XX PR 17-AUG-1999; 99US-0149633P.

CC method for modulating blood cell proliferation in a patent. A claimed
 CC cytokine-inducing composition comprises a pharmaceutical carrier and an
 CC active agent such as the present peptide. Cytokines induced by this
 CC peptide in human leucocyte cultures include interferon-gamma, tumour
 CC necrosis factor-alpha, interleukin-6 and interleukin-10. It was one of
 CC the best overall inducers in almost all cytokine and blood cell
 CC proliferation experiments conducted

XX Sequence 18 AA;

Query Match 100.0%; Score 98; DB 5; Length 18;
 Best Local Similarity 100.0%; Pred. NO. 3.7e-07;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DQPPDVEKPDLPQFQVQS 18
 |||||
 Db 1 DQPPDVEKPDLPQFQVQS 18

RESULT 7

AAE20230
 ID AAE20230 standard; peptide; 18 AA.

XX AAE20230;

DT 18-JUN-2002 (first entry)

DE Colostatinin constituent peptide #3.

XX Blood cell regulator; colostatinin; constituent peptide; oxidative stress;
 KW therapy; oxidative damage; skin; aging; wound healing; cell replacement;
 KW tissue; organ; cosmetic procedure; repair; regeneration; preservation;
 KW transplantation; implantation; dermatological; vulnery.

XX Unidentified.

Key Location/Qualifiers
 FH Modified-site 18
 FT /note= "Optionally C-terminal amide"

PN WO200213850-A1.

XX 21-FEB-2002.

PD 17-AUG-2000; 2000WO-US022776.

PF 17-AUG-2000; 2000WO-US022776.

PR (TEXA) UNIV TEXAS SYSTEM.

XX Stanton GJ, Hughes TK, Boldogh I;

PI WPI; 2002-269151/31.

XX Composition useful for the modulation of blood cell proliferation in a
 PT patient comprises a blood cell regulator selected from colostatinin, its
 FT constituent peptide and/or analog.

PS Claim 6; Page 25; 51pp; English.

XX The invention relates to a composition which comprises a blood cell
 CC regulator selected from colostatinin, its constituent peptide and/or
 CC analogue. The invention is used for modulating the oxidative stress level
 CC in a cell e.g. mammalian or human cell present in a cell culture, tissue,
 CC organ, or organism; or for treating oxidative damage to the skin of a
 CC patient e.g. animal or human; to modulate oxidative stress during/ after
 CC a premature birth or normal birth, preventing/delaying aging in a
 CC patient, enhancing wound healing, and the reduction of side effects of
 CC cosmetic procedures. The method changes the level of an oxidizing species
 CC in the cell, such as decreases or prevents increase in the level of
 CC damage to a biomolecule of the patient selected from DNA, protein and/or
 CC lipid, compared to the same conditions when the oxidative stress
 CC regulator is not present. The modulation of oxidative stress results in

CC enhanced repair, regeneration, and replacement of cells, tissues and
 CC organs (e.g. kidney, liver, pancreas, skin, and the other internal and
 CC external organs), as well as enhanced preservation of such organs for
 CC transplantation, implantation, or scientific research. The present
 CC sequence is a colostatinin constituent peptide

SQ Sequence 18 AA;

Query Match 100.0%; Score 98; DB 5; Length 18;
 Best Local Similarity 100.0%; Pred. NO. 3.7e-07;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DQPPDVEKPDLPQFQVQS 18
 |||||
 Db 1 DQPPDVEKPDLPQFQVQS 18

RESULT 8

ADN60297
 ID ADN60297 standard; peptide; 18 AA.

XX ADN60297;

DT 29-JUL-2004 (first entry)

DE Constituent peptide of colostatinin SEQ ID NO:3.

XX modulator; colostatinin; intracellular signaling molecule modulator;
 KW 4HNE-mediated lipid peroxidation down regulation; apoptosis inhibition;
 KW DNA damage; beta-amyloid; retinoic acid; cytostatic; 4HNE inhibitor;
 KW 4HNE-protein adduct formation reduction;
 KW 4HNE-mediated glutathione depletion inhibition; p53 inhibition;
 KW c-Jun NH2-terminal kinase inhibition.

XX Synthetic.

OS WO2004037851-A2.

PN 06-MAY-2004.

PD 22-OCT-2003; 2003WO-US033423.

PF 22-OCT-2002; 2002US-0420369P.

PR (TEXA) UNIV TEXAS SYSTEM.

XX (BOLD/) BOLDOGH I.

XX (STAN/) STANTON J G.

XX (GEOR/) GEORGIADIS J A.

XX (HUGH/) HUGHES T K.

XX (KRUZ/) KRUZEL M.

PI Boldogh I, Stanton JG, Georgiades JA, Hughes TK, Kruzel M;

WPI; 2004-365494/34.

XX Use of colostatinin for e.g. modulating an intracellular signaling

PT molecule in a cell, down regulating 4HNE-mediated lipid peroxidation in a

FT cell, inhibiting apoptosis in a cell, or protecting against DNA damage in

PT a cell.

PS Claim 6; SEQ ID NO 3; 46pp; English.

XX The present invention describes the use of a modulator selected from
 CC colostatinin, its constituent peptide, its active analogue, and a
 CC combination of these, for modulating an intracellular signaling molecule
 CC in a cell, down regulating 4HNE-mediated lipid peroxidation in a cell,
 CC inhibiting apoptosis in a cell, protecting against DNA damage in a cell,
 CC and reducing the toxic effect of beta-amyloid or retinoic acid on a cell.
 CC The modulator has cytostatic activity, and can be used as a 4HNE
 CC inhibitor. The modulator is useful in the manufacture of a medicament for
 CC reducing 4HNE-protein adduct formation, inhibiting 4HNE-mediated
 CC glutathione depletion, inhibiting 4HNE-induced activation of p53 protein,
 CC and/or inhibiting 4HNE-induced activation of c-Jun NH2-terminal kinases.

CC Colostrinin, or its constituent peptide or active analogue is useful for
 CC inhibiting apoptosis in a cell, protecting against DNA damage in a cell,
 CC or reducing the toxic effect of beta-amyloid or retinoic acid on a cell.
 CC The present sequence represents a synthetic constituent peptide of
 CC colostrinin, which can be used as a modulator in the present invention.
 XX
 XX
 SQ Sequence 18 AA;

Query Match 100.0%; Score 98; DB 8; Length 18;
 Best Local Similarity 100.0%; Pred. No. 3.7e-07;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DQPPDVEKPDLPFFQVQS 18
 |||||
 Db 1 DQPPDVEKPDLPFFQVQS 18
 |||||

RESULT 9
 ADS74401
 ID ADS74401 standard; peptide; 18 AA.

AC ADS74401;

DT 16-DEC-2004 (first entry)

DS Ovine colostrinin peptide.

KW Colostrum; colostrinin; sheep; peptide purification.

XX OS

PN WO2004081038-A1.

XX PD 23-SEP-2004.

XX PF 10-MAR-2004; 2004WO-GB001014.

XX PR 11-MAR-2003; 2003GB-00005552.

XX PR 08-MAR-2004; 2004GB-00005190.

XX PA (REG-) REGEN THERAPEUTICS PLC.

XX PI Georgiades JA, Polanowski A, Wilusz T, Kruzel ML;

XX DR WPI; 2004-677519/66.

XX Recovering peptides such as colostrinin from mammalian colostrum, by
 PT mixing colostrum with alcohol to form alcohol phase containing peptides
 PT and precipitate, separating alcohol phase from precipitate, and
 PT recovering alcohol phase.

XX PS Disclosure, SEQ ID NO 18; 41pp; English.

XX The present sequence is that of a peptide that can be recovered from
 CC ovine colostrinin using the method of the invention. The invention
 CC provides a method for the recovery of peptides (especially colostrinin)
 CC from colostrum in substantially pure, biologically active form and in
 CC high yield. The method involves mixing the colostrum with an alcohol to
 CC form an alcohol phase containing the colostrinin and a precipitate
 CC containing higher molecular weight caseins and other proteins. Best
 CC results are obtained using methanol or ethanol of at least 80%, and
 CC preferably up to 100% purity. The alcohol phase is then separated from
 CC the precipitate, and the colostrinin is separated from the alcohol, is
 CC preferably by evaporation, to form a colostrinin-rich phase, which is
 CC recovered. A precipitation agent, such as ammonium sulfate, may be added
 CC either to the alcohol phase or, preferably, to the colostrinin-rich phase
 CC to induce precipitation of the colostrinin peptides. The method is
 CC generally applicable to the separation of peptides from fluids containing
 CC higher molecular weight proteins, lipids, carbohydrates and/or nucleic
 CC acids.

XX SQ Sequence 18 AA;

Query Match 100.0%; Score 98; DB 8; Length 18;
 Best Local Similarity 100.0%; Pred. No. 3.7e-07;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DQPPDVEKPDLPFFQVQS 18
 |||||
 Db 1 DQPPDVEKPDLPFFQVQS 18
 |||||

RESULT 10
 AAB59355
 ID AAB59355 standard; peptide; 19 AA.

XX AC AAB59355;

XX DT 21-MAR-2001 (first entry)

XX DE Ewe colostrinin peptide fragment derived sequence #15.

XX KW Sheep; colostrinin; proline rich polypeptide; colostrum; immune disorder;
 KW central nervous system disorder; dietary supplement; beta-amyloid plaque.

XX OS Ovis sp.

XX PN WO200075173-A2.

XX PD 14-DEC-2000.

XX PF 02-JUN-2000; 2000WO-GB002128.

XX PR 02-JUN-1999; 99GB-00012852.

XX PA (REG-) REGEN THERAPEUTICS PLC.

XX PI Georgiades JA;

XX DR WPI; 2001-071058/08.

XX PT Peptides having an N-terminal amino acid sequence isolated from
 PT colostrinin for treating e.g. disorders of the central nervous system and
 PT immune system, viral and bacterial infections, and diseases characterized
 PT by amyloid plaques.

XX PS Claim 8; Page 27; 63pp; English.

XX CC The present invention provides the sequences of a number of peptides
 CC found in ewe's colostrinin. Colostrinin is the proline-rich polypeptide
 CC fragment of colostrum. These peptides can be used in the treatment of
 CC central nervous system disorders such as senile dementia, Parkinson's
 CC disease, Alzheimer's disease, psychosis and neurosis, immune system
 CC disorders such as bacterial and viral infections, to improve the
 CC development of a child's immune system, as a dietary supplement, and to
 CC promote the dissolution of beta-amyloid plaques

XX SQ Sequence 19 AA;

Query Match 100.0%; Score 98; DB 4; Length 19;
 Best Local Similarity 100.0%; Pred. No. 3.9e-07;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DQPPDVEKPDLPFFQVQS 18
 |||||
 Db 2 DQPPDVEKPDLPFFQVQS 19
 |||||

RESULT 11
 ABU28927

ID ABU28927 standard; protein; 1047 AA.

XX AC ABU28927;

XX DT 19-JUN-2003 (first entry)

XX

osteoporosis; neurological disease; Alzheimer's disease;
 Parkinson's disease; dementia; short memory; cancer;
 sense or motor function; emotional reaction; fear response; panic;
 osteopathic; neuroprotective; nootropic; antiparkinsonian; cytosstatic;
 tranquilliser.
 KW OS Homo sapiens.
 XX OS EP1447413-A2.
 XX PD 18-AUG-2004.
 XX PF 12-FEB-2004; 2004EP-00003145.
 XX PR 14-FEB-2003; 2003JP-00102207.
 XX PR 09-MAY-2003; 2003JP-00131452.
 XX PA (REAS-) RES ASSOC BIOTECHNOLOGY.
 XX XX Isogai T, Yamamoto J, Nishikawa T, Isono Y, Sugiyama T, Otsuki T;
 PI Wakamatsu A, Ishii S, Nagai K, Irie R;
 PI WPI; 2004-583265/57.
 DR N-PSDB; ADR06500.
 XX XX New 1995 cDNA, useful for treating osteoporosis, neurological diseases,
 PT Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.
 XX XX Claim 1; SEQ ID NO 1962; 2686pp; English.
 XX XX This invention relates to novel, isolated full length human cDNA
 CC molecules and the encoded proteins thereof. Specifically, it refers to
 CC cDNA clones obtained by an oligo-capping method, where none of these
 CC clones are identical to any known human mRNAs. The present invention
 CC describes an immunoassay to identify agonists and antagonists, as well as
 CC antibodies, antisense molecules and siRNAs that can all be used to bind
 CC to and modulate expression of the cDNA molecules. As such, these
 CC molecules are useful for diagnostic markers or therapeutic targets for
 CC the various diseases or morbid states. In particular, they are useful in
 CC gene therapy for treating osteoporosis, neurological disease, Alzheimer's
 CC disease, Parkinson's disease, dementia, short memory and various cancers,
 CC as well as for maintaining equilibrium of sense or motor function, and
 CC for treating emotional reaction, fear response and panic. Accordingly,
 CC they exhibit osteopathic, neuroprotective, nootropic, antiparkinsonian,
 CC cytosstatic and tranquilliser activities. This polypeptide is a protein
 CC encoded by a full length human cDNA sequence of the invention. NOTE: This
 CC sequence is not given in the sequence listing of the specification but
 CC can be obtained on CD-ROM from the European Patent Office, Vienna Sub-
 CC office.
 XX SQ Sequence 116 AA;
 Query Match 49.0%; Score 48; DB 8; Length 116;
 Best Local Similarity 72.7%; Pred. NO. 37;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 Qy 3 PPDEVEKPDLP 13
 Db 100 PPDEVEKPDLP 110
 RESULT 14
 ADJ71952
 ID ADJ71952 standard; protein; 329 AA.
 XX AC ADJ71952;
 XX XX 06-MAY-2004 (first entry)
 XX DE Human PMMW protein amino acid sequence SeqID12.
 XX KW protein modification and maintenance molecule; PMMW; cytosstatic;
 KW antiarteriosclerotic; anticonvulsant; nootropic; neuroprotective;

cerebroprotective; vasotropic; anti-HIV; antiallergic; antimicrobial;
 antiinflammatory; endocrine-gen; thyromimetic;
 cell proliferative disorder; cancer; atherosclerosis;
 neurological disorder; epilepsy; Huntington's disease; stroke;
 immune disorder; inflammatory disorder; AIDS; allergy;
 developmental disorder; Hypothyroidism; Cushing's syndrome; infection;
 human.
 KW OS Homo sapiens.
 XX XX WO2004009797-A2.
 XX PD 29-JAN-2004.
 XX PF 23-JUL-2003; 2003WO-US023249.
 XX PR 23-JUL-2002; 2002US-0398143P.
 PR 09-AUG-2002; 2002US-0402458P.
 PR 12-AUG-2002; 2002US-0403289P.
 PR 27-AUG-2002; 2002US-0406472P.
 PR 06-SEP-2002; 2002US-0409354P.
 XX (INCY-) INCYTE CORP.
 PA Elliott VS, Swarnakar A, Griffin JA, Lee EA, Sprague WW;
 PI Hafalia AJA, Lee SY, Kable AE, Ison CH, Khare R, Chawla NK;
 PI Marquis JP, Jiang X, Jackson AA, Becha SD, Emerling BM, Jin P;
 PI Wilson AD, Richardson TW, Yang J, Baughn MR, Gandhi AR, Nguyen DB;
 PI Ramkumar J, Kallick DA, Kearney L, Lu DM, Gietzen KJ, Tribouley CM;
 PI Lal PG, Blake JJ, Lu Y, Arvizu CS;
 XX WPI; 2004-123392/12.
 DR N-PSDB; ADJ72014.
 XX XX New human protein modification and maintenance molecules (PMMW), useful
 PT for diagnosing, treating and preventing diseases or conditions associated
 PT with the aberrant PMMW expression e.g. cancer, AIDS, epilepsy, or
 PT infections.
 XX Claim 1; SEQ ID NO 12; 387pp; English.
 XX XX This invention relates to novel protein modification and maintenance
 CC molecules (PMMW) and polynucleotides which identify and encode PMMW. The
 CC invention may be useful for the development of compositions with a
 CC cytosstatic, antiarteriosclerotic, anticonvulsant, nootropic,
 CC neuroprotective, cerebroprotective, vasotropic, anti-HIV, antiallergic,
 CC antimicrobial, antinflammatory, endocrine-gen or thyromimetic activity.
 CC The polypeptides and polynucleotides are useful in diagnosing, treating
 CC and preventing diseases or conditions associated with the decreased
 CC expression or overexpression of PMMW, such as cell proliferative (for
 CC example cancer, atherosclerosis), neurological (for example epilepsy,
 CC Huntington's disease, stroke), immune/inflammatory (for example AIDS,
 CC allergies) and developmental (for example Hypothyroidism, Cushing's
 CC syndrome) disorders, or infections. These are also useful in assessing
 CC the effects of exogenous compounds on the expression of nucleic acid and
 CC amino acid sequences of PMMW. The present sequence is that of a human
 CC PMMW protein of the invention.
 XX SQ Sequence 329 AA;
 Query Match 49.0%; Score 48; DB 8; Length 329;
 Best Local Similarity 72.7%; Pred. No. 1.1e+02;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 Qy 3 PPDEVEKPDLP 13
 Db 313 PPDEVEKPDLP 323
 RESULT 15
 AAM79068
 ID AAM79068 standard; protein; 377 AA.
 XX

AC AAM79068;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human protein SEQ ID NO 1730.
 XX
 KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorder; arthritis; inflammation.
 XX
 OS Homo sapiens.
 XX
 PN WO200157190-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 05-FEB-2001; 2001WO-US004098.
 XX
 PR 03-FEB-2000; 2000US-00496914.
 PR 27-APR-2000; 2000US-00560875.
 PR 20-JUN-2000; 2000US-00598075.
 PR 19-JUL-2000; 2000US-00620325.
 PR 01-SEP-2000; 2000US-00654936.
 PR 15-SEP-2000; 2000US-00663561.
 PR 20-OCT-2000; 2000US-00693325.
 PR 30-NOV-2000; 2000US-00728422.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;
 PI Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
 PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
 XX
 DR WPI; 2001-476283/51.
 DR N-PSDB; AAK52201.
 XX
 PT Nucleic acids encoding polypeptides with cytokine-like activities, useful
 PT in diagnosis and gene therapy.
 XX
 PS Claim 20; Page 4066-4067; 6221pp; English.
 XX
 CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
 CC encoded polypeptides (AAM78323-AAK80302) that exhibit activity relating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
 CC (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the
 CC sequence listing were missing at the time of publication
 XX
 SQ Sequence 377 AA;
 Query Match 49.0%; Score 48; DB 4; Length 377;
 Best Local Similarity 72.7%; Pred. No. 1.2e+02;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 Qy 3 PPVKEPDLQP 13
 Db 361 PPVKEPDLQP 371

Search completed: April 12, 2005, 08:18:17
 Job time : 102.8 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 12, 2005, 07:49:30 ; Search time 21.1091 Seconds
(without alignments)
63.654 Million cell updates/sec

Title: US-10-691-330-3

Perfect score: 98

Sequence: 1 DQPPDVEKPDLPQFQVQS 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*

- 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	98	100.0	18	4	US-09-641-803-3
2	51	52.0	1056	4	US-09-134-000C-5086
3	48	49.0	400	4	US-09-134-000C-4389
4	46	46.9	180	4	US-09-744-128-17
5	46	46.9	194	4	US-09-744-128-16
6	45	45.9	26	3	US-09-406-781-13
7	45	45.9	26	4	US-09-880-132-13
8	44.5	45.4	118	4	US-09-270-767-37365
9	44.5	45.4	118	4	US-09-270-767-52582
10	43.5	44.4	593	4	US-09-248-796A-15435
11	43	43.9	1462	3	US-07-792-600-31
12	43	43.9	1462	3	US-09-157-021-31
13	43	43.9	1462	3	US-09-156-842-31
14	43	43.9	1462	4	US-09-591-514-31
15	42	42.9	457	4	US-09-631-603-23
16	42	42.9	457	4	US-09-694-519-1
17	42	42.9	457	4	US-09-826-509-579
18	42	42.9	460	1	US-08-630-592-7
19	42	42.9	460	1	US-08-714-991-7
20	42	42.9	460	3	US-09-032-365A-8
21	42	42.9	460	4	US-09-694-519-2
22	42	42.9	506	1	US-08-631-200-8
23	42	42.9	506	2	US-08-829-553-8
24	42	42.9	506	2	US-08-922-267A-8
25	42	42.9	506	2	US-08-936-707A-8
26	42	42.9	506	2	US-08-936-706A-8
27	42	42.9	506	3	US-09-248-203-8

28	42	42.9	506	3	US-09-032-365A-62
29	42	42.9	506	3	US-08-812-824-4
30	42	42.9	506	3	US-09-406-071-8
31	42	42.9	506	3	US-08-955-918C-10
32	42	42.9	506	3	US-08-697-766A-10
33	42	42.9	506	4	US-09-814-986-8
34	42	42.9	512	3	US-09-032-365A-60
35	42	42.9	518	3	US-09-032-365A-58
36	42	42.9	561	1	US-08-714-991-27
37	42	42.9	561	3	US-09-032-365A-10
38	42	42.9	604	4	US-09-489-039A-8519
39	41.5	42.3	241	4	US-09-248-796A-23201
40	41.5	42.3	553	4	US-09-252-991A-22589
41	41.5	42.3	721	4	US-09-949-016-9402
42	41	41.8	27	2	US-08-563-892A-2
43	41	41.8	27	2	US-08-563-892A-23
44	41	41.8	27	2	US-08-563-892A-24
45	41	41.8	27	2	US-08-563-892A-25

ALIGNMENTS

RESULT 1

US-09-641-803-3

; Sequence 3, Application US/09641803

; Patent No. 6500798

; GENERAL INFORMATION:

; APPLICANT: STANTON, G. John

; APPLICANT: HUGHES, Thomas K.

; APPLICANT: BOLDOGH, Istvan

; TITLE OF INVENTION: USE OF COLOSTRININ, CONSTITUENT PEPTIDES THEREOF, AND

; TITLE OF INVENTION: ANALOGS THEREOF AS OXIDATIVE STRESS REGULATORS

; FILE REFERENCE: 265.00220101

; CURRENT APPLICATION NUMBER: US/09/641.803

; PRIOR FILING DATE: 2000-08-17

; PRIOR FILING DATE: 1999-08-17

; NUMBER OF SEQ ID NOS: 34

; SOFTWARE: Patentin Ver. 2.1

; SEQ ID NO 3

; LENGTH: 18

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: synthetic

; OTHER INFORMATION: peptide

US-09-641-803-3

Query Match 100.0%; Score 98; DB 4; Length 18;

Best Local Similarity 100.0%; Pred. No. 1.4e-08;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DQPPDVEKPDLPQFQVQS 18

|||||

Db 1 DQPPDVEKPDLPQFQVQS 18

RESULT 2

US-09-134-000C-5086

; Sequence 5086, Application US/09134000C

; Patent No. 6617156

; GENERAL INFORMATION:

; APPLICANT: Lynn Doucette-Stamm et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO

; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 032796-032

; CURRENT APPLICATION NUMBER: US/09/134.000C

; CURRENT FILING DATE: 1998-08-13

; PRIOR FILING DATE: 1997-08-15

; NUMBER OF SEQ ID NOS: 6812

; SOFTWARE: Patentin version 3.1

; SEQ ID NO 5086
; LENGTH: 1056
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-5086

Query Match 52.0%; Score 51; DB 4; Length 1056;
Best Local Similarity 47.1%; Pred. No. 16;
Matches 8; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 1 DQPDVEKPDLPQFQV 17
DB 265 DQVDLQKETKQFOLK 281

RESULT 3

US-09-134-000C-4389
; Sequence 4389, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4389
; LENGTH: 400
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-4389

Query Match 49.0%; Score 48; DB 4; Length 400;
Best Local Similarity 62.5%; Pred. No. 16;
Matches 10; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 DQPDVEKPDLPQFQV 16
DB 133 DQFGKEKWDLPQMEV 148

RESULT 4

US-09-744-128-17
; Sequence 17, Application US/09744128
; Patent No. 6677306
; GENERAL INFORMATION:
; APPLICANT: Veis et al.
; TITLE OF INVENTION: Chondrogenic and Osteogenic Inducing Molecule
; FILE REFERENCE: 27636/36983
; CURRENT APPLICATION NUMBER: US/09/744,128
; CURRENT FILING DATE: 2001-05-16
; PRIOR FILING DATE: PCT/US99/17342
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: 60/094,489
; PRIOR FILING DATE: 1998-07-29
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn 3.1
; SEQ ID NO 17
; LENGTH: 180
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial sequence: PCR product
US-09-744-128-17

Query Match 46.9%; Score 46; DB 4; Length 180;
Best Local Similarity 50.0%; Pred. No. 13;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 3 PPDVKEPDLPQFQV 18
DB 105 PPSAQPFQPFQFQA 120

RESULT 5

US-09-744-128-16
; Sequence 16, Application US/09744128
; Patent No. 6677306
; GENERAL INFORMATION:
; APPLICANT: Veis et al.
; TITLE OF INVENTION: Chondrogenic and Osteogenic Inducing Molecule
; FILE REFERENCE: 27636/36983
; CURRENT APPLICATION NUMBER: US/09/744,128
; CURRENT FILING DATE: 2001-05-16
; PRIOR FILING DATE: PCT/US99/17342
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: 60/094,489
; PRIOR FILING DATE: 1998-07-29
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn 3.1
; SEQ ID NO 16
; LENGTH: 194
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial sequence: PCR product
US-09-744-128-16

Query Match 46.9%; Score 46; DB 4; Length 194;
Best Local Similarity 50.0%; Pred. No. 14;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 3 PPDVKEPDLPQFQV 18
DB 119 PPSAQPFQPFQFQA 134

RESULT 6

US-09-406-781-13
; Sequence 13, Application US/09406781
; Patent No. 6306663
; GENERAL INFORMATION:
; APPLICANT: Kenten, John
; APPLICANT: Roberts, Steven
; TITLE OF INVENTION: CONTROLLING PROTEIN LEVELS IN EUKARYOTIC ORGANISMS
; FILE REFERENCE: 2757-3
; CURRENT APPLICATION NUMBER: US/09/406,781
; CURRENT FILING DATE: 1999-09-28
; EARLIER FILING DATE: 60/119,851
; EARLIER FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: PEST example
; OTHER INFORMATION: sequence
US-09-406-781-13

Query Match 45.9%; Score 45; DB 3; Length 26;
Best Local Similarity 57.1%; Pred. No. 2.2;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 3 PPDVKEPDLPQFQV 16
DB 2 PPGVEEDVGPLPV 15

RESULT 7

US-09-880-132-13

; Sequence 13, Application US/09880132
; Patent No. 655280
; GENERAL INFORMATION:
; APPLICANT: Kenten, John
; APPLICANT: Roberts, Steven
; TITLE OF INVENTION: CONTROLLING PROTEIN LEVELS IN EUKARYOTIC ORGANISMS
; FILE REFERENCE: 2757-6
; CURRENT APPLICATION NUMBER: US/09/880,132
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: 09/406,781
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: 60/119,851
; PRIOR FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: PEST example
; OTHER INFORMATION: sequence
US-09-880-132-13

Query Match 45.9%; Score 45; DB 4; Length 26;
Best Local Similarity 57.1%; Pred. No. 2.2;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 3 PPDEVKPDLPQFQV 16
|||:||||
Db 2 PPGVEEPDVGLPV 15

RESULT 8
US-09-270-767-37365
; Sequence 37365, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 37365
; LENGTH: 118
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-37365

Query Match 45.4%; Score 44.5; DB 4; Length 118;
Best Local Similarity 66.7%; Pred. No. 14;
Matches 10; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

Qy 2 QPPDEVKPDLPQFQ 15
|||:||||
Db 23 QPVDITKADLKQPFQ 37

RESULT 9
US-09-270-767-52582
; Sequence 52582, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 52582

; LENGTH: 118
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-52582

Query Match 45.4%; Score 44.5; DB 4; Length 118;
Best Local Similarity 66.7%; Pred. No. 14;
Matches 10; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

Qy 2 QPPDEVKPDLPQFQ 15
|||:||||
Db 23 QPVDITKADLKQPFQ 37

RESULT 10
US-09-248-796A-15435
; Sequence 15435, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 15435
; LENGTH: 593
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-15435

Query Match 44.4%; Score 43.5; DB 4; Length 593;
Best Local Similarity 58.8%; Pred. No. 1.2e+02;
Matches 10; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

Qy 2 QPPDEVKPDLPQFQVS 18
|||:||||
Db 363 QPPQLEPPVLEP-QVSS 378

RESULT 11
US-07-792-600-31
; Sequence 31, Application US/07792600
; Patent No. 6008045
; GENERAL INFORMATION:
; APPLICANT: COPELAND, WILLIAM C.
; APPLICANT: WANG, TERESA S.-F.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: TEMPLATE-DEPENDENT ENZYMATIC SYNTHESIS OF NUCLEIC ACID
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Peter G. Carroll
; STREET: 220 Montgomery Street, Suite 710
; CITY: San Francisco
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/792,600
; FILING DATE: 19911115
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: CARROLL, PETER G.

```

RESULT 13
US-09-156-842-31
  : Sequence 31, Application 'US/09156842A
  : Patent No. 6103473
  : GENERAL INFORMATION:
  : APPLICANT: Copeland, William C.
  : APPLICANT: Wang, Teresa S. F.
  : TITLE OF INVENTION: Drug Screening
  : FILE REFERENCE: STDU-03485
  : CURRENT APPLICATION NUMBER: US/09/156,842A
  : CURRENT FILING DATE: 1998-09-18
  : EARLIER APPLICATION NUMBER: 07/792,600
  : EARLIER FILING DATE: 1991-11-15
  : NUMBER OF SEQ ID NOS: 35
  : SOFTWARE: Patentin Ver. 2.0
  : SEQ ID NO 31
  : LENGTH: 1462
  : TYPE: PRT
  : ORGANISM: Homo sapiens
US-09-156-842-31

```

Qy 2 QPPDVEKPDLPF 14
:|:|:|:
Db 324 RPPDIRKSDSSPY 336

Search completed: April 12, 2005, 08:07:50
Job time : 22.1091 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 12, 2005, 07:49:30 ; Search time 55.8 Seconds
(without alignments)
107.096 Million cell updates/sec

Title: US-10-691-330-3

Perfect score: 98

Sequence: 1 DQPDVEKPDLPQFQVQS 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1418010 seqs, 331997259 residues

Total number of hits satisfying chosen parameters: 1418010

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%

Listing first 100%

Database : Published Applications AA:*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	98	100.0	18	14	US-10-281-652-3
2	98	100.0	18	17	US-10-691-330-3
3	51	52.0	1047	15	US-10-282-122A-56851
4	48	49.0	377	15	US-10-210-130-118
5	48	49.0	377	15	US-10-250-613-1
6	48	49.0	541	9	US-09-815-242-10746
7	47	48.0	162	16	US-10-437-963-204178
8	46	45.9	212	15	US-10-437-963-200861
9	46	45.9	313	15	US-10-425-114-52231
10	46	46.9	336	15	US-10-424-599-233789
11	46	46.9	337	15	US-10-425-114-39945
12	45	45.9	26	9	US-09-880-149-13
13	45	45.9	26	9	US-09-880-132-13

14	45	45.9	26	14	US-10-345-281-13	Sequence 13, Appli
15	45	45.9	159	15	US-10-424-599-228647	Sequence 228647,
16	45	45.9	848	9	US-09-925-302-766	Sequence 766, App
17	45	45.9	848	10	US-09-925-302-766	Sequence 766, App
18	45	45.9	1006	16	US-10-437-963-136358	Sequence 136358,
19	45	45.9	1289	16	US-10-437-963-157119	Sequence 157119,
20	44.5	45.4	583	16	US-10-437-963-114326	Sequence 114326,
21	44.5	45.4	852	10	US-09-971-101A-4	Sequence 4, Appli
22	44	44.9	167	17	US-10-485-037-4	Sequence 4, Appli
23	44	44.9	212	9	US-09-925-300-1577	Sequence 1577, Ap
24	44	44.9	251	15	US-10-425-114-43396	Sequence 43396, A
25	43.5	44.4	1367	15	US-10-320-797-3355	Sequence 3355, A
26	43	43.9	52	9	US-09-864-761-39967	Sequence 39967, A
27	43	43.9	300	15	US-10-369-493-12997	Sequence 12997, A
28	43	43.9	387	16	US-10-437-963-118548	Sequence 118548,
29	43	43.9	448	15	US-10-425-114-54044	Sequence 54044, A
30	43	43.9	712	15	US-10-369-493-3977	Sequence 3977, Ap
31	43	43.9	1165	16	US-10-408-765A-1392	Sequence 1392, Ap
32	43	43.9	1212	15	US-10-618-581-5	Sequence 5, Appli
33	42	42.9	51	15	US-10-424-599-247293	Sequence 247293,
34	42	42.9	264	14	US-10-017-161-706	Sequence 706, App
35	42	42.9	340	15	US-10-389-566-1425	Sequence 1425, Ap
36	42	42.9	340	15	US-10-389-566-1831	Sequence 1831, Ap
37	42	42.9	364	15	US-10-425-114-41207	Sequence 41207, A
38	42	42.9	399	9	US-09-764-870-409	Sequence 409, App
39	42	42.9	399	14	US-10-125-540-409	Sequence 409, App
40	42	42.9	421	16	US-10-437-963-133514	Sequence 133514,
41	42	42.9	457	10	US-09-826-509-579	Sequence 579, App
42	42	42.9	457	14	US-10-225-567A-469	Sequence 469, App
43	42	42.9	457	15	US-10-292-798-618	Sequence 618, App
44	42	42.9	457	16	US-10-831-393-1	Sequence 1, Appli
45	42	42.9	457	17	US-10-925-095-579	Sequence 579, App

ALIGNMENTS

RESULT 1

US-10-281-652-3
; Sequence 3, Application US/10281652
; Publication No. US20030091606A1
; GENERAL INFORMATION:
; APPLICANT: STANTON, G. John
; APPLICANT: HUGHES, Thomas K.
; APPLICANT: BOLDOGH, Istvan
; TITLE OF INVENTION: USE OF COLOSTRININ, CONSTITUENT PEPTIDES THEREOF, AND
; FILE REFERENCE: 265.00220101
; CURRENT APPLICATION NUMBER: US/10/281.652
; PRIOR FILING DATE: 2002-10-28
; PRIOR FILING DATE: 2000-08-17
; PRIOR FILING DATE: 1999-08-17
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 3
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: peptide
US-10-281-652-3

Query Match 100.0%; Score 98; DB 14; Length 18;
Best Local Similarity 100.0%; Pred. No. 3.3e-07;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DQPDVEKPDLPQFQVQS 18

Db 1 DQPDVEKPDLPQFQVQS 18

```
RESULT 2
US-10-691-330-3
; Sequence 3, Application US/10691330
; Publication No. US20050042300A1
; GENERAL INFORMATION:
; APPLICANT: Boldogh, Istvan
; APPLICANT: Stanton, G. John
; APPLICANT: Georgiades, Jerzy A.
; APPLICANT: Hughes, Thomas K., Jr.
; APPLICANT: Kruzel, Marian
; TITLE OF INVENTION: USE OF COLOSSTRININ, CONSTITUENT PEPTIDES THEREOF, AND ANALOGS
; TITLE OF INVENTION: THEROF AS INHIBITORS OF APOPTOSIS AND OTHER CELLULAR DAMAGE
; FILE REFERENCE: 265 00390101
; CURRENT APPLICATION NUMBER: US/10/691,330
; CURRENT FILING DATE: 2003-10-22
; PRIOR APPLICATION NUMBER: US 60/420,369
; PRIOR FILING DATE: 2002-10-22
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3
; LENGTH: 18
; TYPE: PRT
; ORGANISM: artificial
; FEATURE:
; OTHER INFORMATION: Synthetic Peptides
US-10-691-330-3

Query Match 100.0%; Score 98; DB 17; Length 18;
Best Local Similarity 100.0%; Pred. No. 3.3e-07;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DQPPDVEKPDLPQFQVQS 18
Db 1 DQPPDVEKPDLPQFQVQS 18

RESULT 3
US-10-282-122A-56851
; Sequence 56851, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931

US-10-691-330-3.rapb
```

```
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 56851
; LENGTH: 1047
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-10-282-122A-56851

Query Match 52.0%; Score 51; DB 15; Length 1047;
Best Local Similarity 47.1%; Pred. No. 1.1e+02;
Matches 8; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 1 DQPPDVEKPDLPQFQVQ 17
Db 256 DQPPDVEKPDLPQFQVQ 272

RESULT 4
US-10-210-130-118
; Sequence 118, Application US/10210130
; Publication No. US20040014053A1
; GENERAL INFORMATION:
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Patturajan, Meera
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Miller, Charles E.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Pena, Carol E.A.
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Li, Li
; APPLICANT: Berghs, Constance
; APPLICANT: Zhong, Mei
; APPLICANT: Casman, Stacie J.
; APPLICANT: Voss, Edward Z.
; APPLICANT: Boldog, Ferenc L.
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Smithson, Glennda
; APPLICANT: Ji, Weizhen
; APPLICANT: Gorman, Linda
; APPLICANT: Vernet, Corine A.M.
; APPLICANT: Leite, Mario W.
; APPLICANT: Guo, Xiaojia Sasha
; APPLICANT: Anderson, David W.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Burgess, Catherine E.
; APPLICANT: Khrantsov, Nikolai V.
; APPLICANT: Ort, Tatiana
; APPLICANT: Ellerman, Karen
; APPLICANT: Rastelli, Luca
; APPLICANT: Agce, Michele L.
; APPLICANT: Chaudhuri, Amitabha
; APPLICANT: Chant, John S.
; APPLICANT: DiPippo, Vincent A.
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Eisen, Andrew J.
; APPLICANT: Gangolli, Esha A.
; APPLICANT: Giot, Loic
; APPLICANT: Ooi, Chean Eng
; APPLICANT: Rothenberg, Mark E.
; APPLICANT: Spaderna, Steven K.
; APPLICANT: Hjal, Tord
; APPLICANT: Liu, Xiaohong
; APPLICANT: Taupier, Raymond J., Jr.
; APPLICANT: Catterton, Elina
; APPLICANT: Shenoy, Suresh G.
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
```

```
FILE REFERENCE: 21402-416C (Cura-716 SWT)
CURRENT APPLICATION NUMBER: US/10/210,130
CURRENT FILING DATE: 2002-08-01
PRIOR APPLICATION NUMBER: 60/309,501
PRIOR FILING DATE: 2001-08-02
PRIOR APPLICATION NUMBER: 60/316,508
PRIOR FILING DATE: 2001-08-31
PRIOR APPLICATION NUMBER: 60/354,655
PRIOR FILING DATE: 2002-02-05
PRIOR APPLICATION NUMBER: 60/310,291
PRIOR FILING DATE: 2001-08-03
PRIOR APPLICATION NUMBER: 60/383,887
PRIOR FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: 60/310,951
PRIOR FILING DATE: 2001-08-08
PRIOR APPLICATION NUMBER: 60/323,936
PRIOR FILING DATE: 2001-09-21
PRIOR APPLICATION NUMBER: 60/381,039
PRIOR FILING DATE: 2002-05-16
PRIOR APPLICATION NUMBER: 60/311,292
PRIOR FILING DATE: 2001-08-09
PRIOR APPLICATION NUMBER: 60/311,979
PRIOR FILING DATE: 2001-08-13
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 369
SOFTWARE: Curaseq version 0.1
SEQ ID NO 118
LENGTH: 377
TYPE: PRT
ORGANISM: Homo sapiens
US-10-210-130-118

Query Match          49.0%; Score 48; DB 15; Length 377;
Best Local Similarity 72.7%; Pred. No. 1e+02;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy      3 PPDVEKPDLPQ 13
Db      361 PPDVEQPTQP 371

RESULT 5
US-10-250-613-1
Sequence 1, Application US/10/250,613
Publication No. US2004098828A1
GENERAL INFORMATION:
APPLICANT: LU, Dyung Aina M.; BAUGHN, Mariah R.;
APPLICANT: YAO, Monique G.; DING, Li;
APPLICANT: HONCHELL, Cynthia D.; YUE, Henry;
APPLICANT: TANG, Y. Tom; WARREN, Bridget A.;
APPLICANT: DUGAN, Brendan M.; XU, Yuning;
APPLICANT: CHAWLA, Narinder K.; GRIFFIN, Jennifer A.;
APPLICANT: STEWART, Elizabeth A.; GANDHI, Ameena R.;
APPLICANT: KHAN, Farrah A.; THANGAVELU, Kavitha;
APPLICANT: ISON, Craig H.; AZIMZAI, Yalda;
APPLICANT: HAPALIA, April J.A.; GIETZEN, Kimberly J.;
APPLICANT: LAL, Preeti G.; SANJANWALA, Madhusudan M.;
APPLICANT: ELLIOTT, Vicki S.
TITLE OF INVENTION: CYTOSKELETAL-ASSOCIATED PROTEINS
FILE REFERENCE: PF-0878 USN
CURRENT APPLICATION NUMBER: US/10/250,613
CURRENT FILING DATE: 2003-07-02
PRIOR APPLICATION NUMBER: PCT/US02/00178
PRIOR FILING DATE: 2002-01-04
PRIOR APPLICATION NUMBER: US 60/260,085
PRIOR FILING DATE: 2001-01-04
PRIOR APPLICATION NUMBER: US 60/268,554
PRIOR FILING DATE: 2001-02-13
PRIOR APPLICATION NUMBER: US 60/269,111
PRIOR FILING DATE: 2001-02-14
PRIOR APPLICATION NUMBER: US 60/271,211
PRIOR FILING DATE: 2001-02-23
NUMBER OF SEQ ID NOS: 36
```

```
SOFTWARE: PERL Program
SEQ ID NO 1
LENGTH: 377
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No: 5566074CD1
US-10-250-613-1

Query Match          49.0%; Score 48; DB 15; Length 377;
Best Local Similarity 72.7%; Pred. No. 1e+02;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy      3 PPDVEKPDLPQ 13
Db      361 PPDVEQPTQP 371

RESULT 6
US-09-815-242-10746
Sequence 10746, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.01A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 10746
LENGTH: 541
TYPE: PRT
ORGANISM: Enterococcus faecalis
US-09-815-242-10746

Query Match          49.0%; Score 48; DB 9; Length 541;
Best Local Similarity 62.5%; Pred. No. 1.5e+02;
Matches 10; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy      1 DQPPDVEKPDLPQV 16
Db      274 DQPGKEKWDLPQMEV 289

RESULT 7
US-10-437-963-204178
Sequence 204178, Application US/10437963
Publication No. US20040123343A1
GENERAL INFORMATION:
```

```
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 204178
; LENGTH: 162
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_99291C.1.pep
US-10-437-963-204178

Query Match      48.0%; Score 47; DB 16; Length 162;
Best Local Similarity 100.0%; Pred. No. 57;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      3 PPDVEKPD 10
Db      15 PPDVEKPD 22

RESULT 8
US-10-437-963-200861
; Sequence 200861, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 200861
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_96290C.1.pep
US-10-437-963-200861

Query Match      46.9%; Score 46; DB 16; Length 212;
Best Local Similarity 52.9%; Pred. No. 1.1e+02;
Matches      9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy      2 QPPDVEKPDLPQFQVQS 18
Db      36 RPPLRSKPEALPFQQAQS 52

RESULT 9
US-10-425-114-52231
; Sequence 52231, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
```

```
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 52231
; LENGTH: 313
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700894460_FLI.pep
US-10-425-114-52231

Query Match      46.9%; Score 46; DB 15; Length 313;
Best Local Similarity 53.3%; Pred. No. 1.6e+02;
Matches      8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy      2 QPPDVEKPDLPQFQV 16
Db      151 QPVELEFPNQPLQV 165

RESULT 10
US-10-424-599-233789
; Sequence 233789, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 233789
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_53136C.1.pep
US-10-424-599-233789

Query Match      46.9%; Score 46; DB 15; Length 336;
Best Local Similarity 53.3%; Pred. No. 1.7e+02;
Matches      8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy      2 QPPDVEKPDLPQFQV 16
Db      149 QPVELEFPNQPLQV 163

RESULT 11
US-10-425-114-39945
; Sequence 39945, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
```


FILE REFERENCE: 38-21(53313)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 39945
LENGTH: 337
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: 700904208_FLI.pcp
US-10-425-114-39945

Query Match 46.9%; Score 46; DB 15; Length 337;
Best Local Similarity 53.3%; Pred. No. 1.7e+02;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 2 PPVVEKPDLPQFQV 16
|||::|||
Db 151 QPVEEPDVGPLPV 165

RESULT 12
US-880-149-13
Sequence 13, Application US/09880149
Patent No. US20020146843A1
GENERAL INFORMATION:
APPLICANT: Kenten, John
TITLE OF INVENTION: CONTROLLING PROTEIN LEVELS IN EUARYOTIC ORGANISMS
FILE REFERENCE: 2757-5
CURRENT APPLICATION NUMBER: US/09/880,149
CURRENT FILING DATE: 2001-06-14
PRIOR APPLICATION NUMBER: 09/406,781
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: 60/119,851
PRIOR FILING DATE: 1999-02-12
NUMBER OF SEQ ID NOS: 67
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 13
LENGTH: 26
TYPE: PRT
ORGANISM: Unknown Organism
FEATURE:
OTHER INFORMATION: Description of Unknown Organism: PEST example
US-09-880-149-13

Query Match 45.9%; Score 45; DB 9; Length 26;
Best Local Similarity 57.1%; Pred. No. 16;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 3 PPVVEKPDLPQFQV 16
|||::|||
Db 2 PPGVEEPDVGPLPV 15

RESULT 13
US-880-132-13
Sequence 13, Application US/09880132
Patent No. US20020173049A1
GENERAL INFORMATION:
APPLICANT: Kenten, John
TITLE OF INVENTION: CONTROLLING PROTEIN LEVELS IN EUARYOTIC ORGANISMS
FILE REFERENCE: 2757-6
CURRENT APPLICATION NUMBER: US/09/880,132
CURRENT FILING DATE: 2001-06-14
PRIOR APPLICATION NUMBER: 09/406,781
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: 60/119,851
PRIOR FILING DATE: 1999-02-12
NUMBER OF SEQ ID NOS: 67
SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 13
LENGTH: 26
TYPE: PRT
ORGANISM: Unknown Organism
FEATURE:
OTHER INFORMATION: Description of Unknown Organism: PEST example
US-09-880-132-13

Query Match 45.9%; Score 45; DB 9; Length 26;
Best Local Similarity 57.1%; Pred. No. 16;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 3 PPVVEKPDLPQFQV 16
|||::|||
Db 2 PPGVEEPDVGPLPV 15

RESULT 14
US-10-345-281-13
Sequence 13, Application US/10345281
Publication No. US20030153727A1
GENERAL INFORMATION:
APPLICANT: Kenten, John
TITLE OF INVENTION: CONTROLLING PROTEIN LEVELS IN EUARYOTIC ORGANISMS
FILE REFERENCE: 2757-6
CURRENT APPLICATION NUMBER: US/10/345,281
CURRENT FILING DATE: 2003-01-16
PRIOR APPLICATION NUMBER: US/09/880,132
PRIOR FILING DATE: 2001-06-14
PRIOR APPLICATION NUMBER: 09/406,781
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: 60/119,851
PRIOR FILING DATE: 1999-02-12
NUMBER OF SEQ ID NOS: 67
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 13
LENGTH: 26
TYPE: PRT
ORGANISM: Unknown Organism
FEATURE:
OTHER INFORMATION: Description of Unknown Organism: PEST example
US-10-345-281-13

Query Match 45.9%; Score 45; DB 14; Length 26;
Best Local Similarity 57.1%; Pred. No. 16;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 3 PPVVEKPDLPQFQV 16
|||::|||
Db 2 PPGVEEPDVGPLPV 15

RESULT 15
US-10-424-599-228647
Sequence 228647, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa Thomas J
APPLICANT: Kovalic David K
APPLICANT: Zhou Yihua
APPLICANT: Cao Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 228647
LENGTH: 159
TYPE: PRT

; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_48496C.1.pap
US-10-424-599-228647

Query Match 45.9%; Score 45; DB 15; Length 159;
Best Local Similarity 58.3%; Pred. No. 1.1e+02;
Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 5 DVEKPDLOPFQV 16
||:||||:|:
Db 145 DVKPDVKEVQI 156

Search completed: April 12, 2005, 07:55:20
Job time : 55.8 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 12, 2005, 07:49:31 ; Search time 85.4182 Seconds
(without alignments)
107.909 Million cell updates/sec

Title: US-10-691-330-3

Perfect score: 98

Sequence: 1 DQPPDVKEKPDLPQPFQVQS 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	55	56.1	483	2 Q9A5S4	Q9A5S4 caulobacter
2	51	52.0	182	2 Q7W511	Q7W511 bordetella
3	51	52.0	182	2 Q7WGH6	Q7WGH6 bordetella
4	51	52.0	182	2 Q9X620	Q9X620 bordetella
5	51	52.0	392	2 Q88168	Q88168 enterococcus
6	51	52.0	1047	2 Q832P0	Q832P0 enterococcus
7	48	49.0	377	1 TTL_HUMAN	Q8ng68 homo sapien
8	48	49.0	461	1 DQOR_BOVIN	P27117 bos taurus
9	48	49.0	541	2 Q835M9	Q835M9 enterococcus
10	48	49.0	652	2 Q8CHL9	Q6chl9 yarrowia li
11	47	48.0	320	2 Q700G4	Q700G4 ustilago ma
12	47	48.0	320	2 Q700G6	Q700G6 ustilago ma
13	47	48.0	320	2 Q700G8	Q700G8 ustilago ma
14	47	48.0	320	2 Q700I9	Q700I9 ustilago ma
15	47	48.0	404	2 Q99118	Q99118 ustilago ma
16	47	48.0	467	2 Q67QH0	Q67qh0 symbiobacte
17	47	48.0	516	2 Q9C8M8	Q9c8m8 arabidopsis
18	47	48.0	719	2 Q9C8M9	Q9c8m9 arabidopsis
19	46.5	47.4	126	2 Q84R17	Q84r17 arabidopsis
20	46	46.9	134	2 Q9X8L8	Q9x8l8 streptomyce
21	46	46.9	137	2 Q9D2D5	Q9d2d5 mus musculu
22	46	46.9	141	2 Q9QW43	Q9qw43 mus sp...a
23	46	46.9	187	2 Q8UEB8	Q8ueb8 agrobacteri
24	46	46.9	199	2 Q7CYE5	Q7cye5 agrobacteri
25	46	46.9	202	2 Q6EQW0	Q6eqw0 oryza sativ
26	46	46.9	210	1 AMEX_MOUSE	P63277 mus musculu
27	46	46.9	210	1 AMEX_RAT	P63278 rattus norv
28	46	46.9	219	2 Q8PCW7	Q8pcw7 mus musculu
29	46	46.9	318	2 Q700K9	Q700k9 ustilago ma
30	46	46.9	320	2 Q700H9	Q700h9 ustilago ma
31	46	46.9	499	2 Q9V539	Q9v539 drosophila

32	46	46.9	627	2 Q7ZXG8	Q7zxg8 xenopus lae
33	46	46.9	669	2 Q6IMN2	Q6imn2 xenopus lae
34	46	46.9	680	2 Q75T15	Q75t15 xenopus lae
35	46	46.9	728	1 SUB2_DEIRA	Q9ym8 deinococcus
36	46	46.9	908	2 Q8MMF4	Q8mmf4 drosophila
37	46	46.9	1424	2 Q7R422	Q7r422 giardia lam
38	45	45.9	194	2 Q88326	Q88326 mesocricatu
39	45	45.9	345	2 Q8C6H1	Q8c6h1 mus musculu
40	45	45.9	369	2 Q73737	Q73737 gallus gall
41	45	45.9	391	1 Y258_HUMAN	Q92546 homo sapien
42	45	45.9	391	2 Q8BH77	Q8bht7 mus musculu
43	45	45.9	415	2 Q6ZQF1	Q6zqf1 mus musculu
44	45	45.9	435	2 Q6YID3	Q6yid3 helobdella
45	45	45.9	570	2 Q6A0A5	Q6a0a5 mus musculu

ALIGNMENTS

RESULT 1

Q9A5S4	PRELIMINARY;	PRT;	483 AA.
AC Q9A5S4	01-JUN-2001 (TrEMBLrel. 17, Created)		
DT 01-JUN-2001	(TrEMBLrel. 17, Last sequence update)		
DE Methylmalonyl-CoA mutase, beta subunit.			
GN OrderedLocusNames=CC2373;			
OS Caulobacter crescentus.			
OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacteriales;			
OX Caulobacteraceae; Caulobacter.			
NCBI_TaxID=155892;			
SEQUENCE FROM N.A.			
STRAIN=ATCC 19089 / CB15;			
MEDLINE=21173698; PubMed=11259647; DOI=10.1073/pnas.061029298;			
Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,			
Eisen J.A., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,			
Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,			
DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,			
Kolony J.F., Smit J., Craven M.B., Khouri H.M., Shetty J.,			
Berry K.J., Utterback T.R., Tran K., Wolf A.M., Vamathevan J.J.,			
Ermolaeva M.D., White O., Salzberg S.L., Venter J.C., Shapiro L.,			
Fraser C.M.;			
"Complete genome sequence of Caulobacter crescentus.";			
Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).			
EMBL; AE005906; AAK24344.1; -			
PIR; D87543; D87543.			
HSSP; P11652; 1REQ.			
TIGR; CC2373; -			
GO; GO:0004494; F:methylmalonyl-CoA mutase activity; IEA.			
GO; GO:0008152; P:metabolism; IEA.			
InterPro; IPR006099; MCOA_mutase.			
Pfam; PF01642; MCOA_mutase; 1.			
Complete proteome.			
SEQUENCE 483 AA; 50032 MW; 194F84D33268D6D5 CRC64;			

Query Match 56.1%; Score 55; DB 2; Length 483;
Best Local Similarity 58.8%; Pred. No. 6.7;
Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy	1 DQPPDVKEKPDLPQPFQVQ 17
Db	439 DKPEVETPDSSAFVQ 455

RESULT 2

Q7W511	PRELIMINARY;	PRT;	182 AA.
AC Q7W511	01-OCT-2003 (TrEMBLrel. 25, Created)		
DT 01-OCT-2003	(TrEMBLrel. 25, Last sequence update)		
DT 01-MAR-2004	(TrEMBLrel. 26, Last annotation update)		

```

DE Outer membrane lipoprotein.
GN Name=OmlA; OrderedLocusNames=BPP3495;
OS Bordetella parapertussis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OX Alcaligenaceae; Bordetella.
RN NCBI_TaxID=519;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=12822 / ATCC BAA-587;
RX MEDLINE=22827954; PubMed=12910271; DOI=10.1038/ng1227;
RA Parkhill J., Sebaihia M., Preston A., Murphy L.D., Thomson N.R.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cerdano-Tarraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Felwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA Rabinowitsch E., Rutter S., Norberczak H., O'Neil S., Ormond D., Price C.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40(2003).
DR EMBL; BX640433; CAE38779.1; -.
DR GO; GO:0019867; C:outer membrane; IEA.
DR InterPro; IPR007450; SmpA_OmlA.
DR Pfam; PF04355; SmpA_OmlA_1.
DR Complete proteome; Lipoprotein.
KW Complete proteome; Lipoprotein.
SQ SEQUENCE 182 AA; 20490 MW; 73F6DB9B1714377F CRC64;

Query Match 52.0%; Score 51; DB 2; Length 182;
Best Local Similarity 72.7%; Pred. No. 9.5;
Matches 8; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 7 EKPDLPFQVQ 17
Db 114 EQPDLPFQIE 124
[1]

RESULT 3
QY 7 EKPDLPFQVQ 17
Db 114 EQPDLPFQIE 124
[1]

ID Q7WGH6 PRELIMINARY; PRT; 182 AA.
AC Q7WGH6;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Outer membrane lipoprotein.
GN Name=OmlA; OrderedLocusNames=BB3943;
OS Bordetella bronchiseptica (Alcaligenes bronchisepticus).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OX Alcaligenaceae; Bordetella.
RN NCBI_TaxID=518;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=RB50 / ATCC BAA-588;
RX MEDLINE=22827954; PubMed=12910271; DOI=10.1038/ng1227;
RA Parkhill J., Sebaihia M., Preston A., Murphy L.D., Thomson N.R.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cerdano-Tarraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Felwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA Rabinowitsch E., Rutter S., Norberczak H., O'Neil S., Ormond D., Price C.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40(2003).
DR EMBL; BX640449; CAE34306.1; -.
DR GO; GO:0019867; C:outer membrane; IEA.
DR InterPro; IPR007450; SmpA_OmlA.
DR Pfam; PF04355; SmpA_OmlA_1.

KW Complete proteome; Lipoprotein.
SQ SEQUENCE 182 AA; 20490 MW; 73F6DB9B1714377F CRC64;

Query Match 52.0%; Score 51; DB 2; Length 182;
Best Local Similarity 72.7%; Pred. No. 9.5;
Matches 8; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 7 EKPDLPFQVQ 17
Db 114 EQPDLPFQIE 124
[1]

RESULT 4
QY 7 EKPDLPFQVQ 17
Db 114 EQPDLPFQIE 124
[1]

ID Q9X6Z0 PRELIMINARY; PRT; 182 AA.
AC Q9X6Z0;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Outer membrane lipoprotein.
GN Name=OmlA; Synonym=OmlA; OrderedLocusNames=BP2508;
OS Bordetella pertussis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OX Alcaligenaceae; Bordetella.
RN NCBI_TaxID=520;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=Tohamai;
RA Pradel E.;
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=Tohamai;
RA Pradel E.;
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=Tohamai I / ATCC BAA-589 / NCTC 13251;
RX MEDLINE=22827954; PubMed=12910271; DOI=10.1038/ng1227;
RA Parkhill J., Sebaihia M., Preston A., Murphy L.D., Thomson N.R.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cerdano-Tarraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Felwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA Rabinowitsch E., Rutter S., Norberczak H., O'Neil S., Ormond D., Price C.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40(2003).
DR EMBL; AJ238308; CAB41013.1; -.
DR EMBL; BX640418; CAB42780.1; -.
DR GO; GO:0019867; C:outer membrane; IEA.
DR InterPro; IPR007450; SmpA_OmlA.
DR Pfam; PF04355; SmpA_OmlA_1.
DR Complete proteome; Lipoprotein.
KW Complete proteome; Lipoprotein.
SQ SEQUENCE 182 AA; 20489 MW; 73F6DB9B171AD791 CRC64;

Query Match 52.0%; Score 51; DB 2; Length 182;
Best Local Similarity 72.7%; Pred. No. 9.5;
Matches 8; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 7 EKPDLPFQVQ 17
Db 114 EQPDLPFQIE 124
[1]

RESULT 5
QY 7 EKPDLPFQVQ 17
Db 114 EQPDLPFQIE 124
[1]

ID O88168 PRELIMINARY; PRT; 392 AA.
AC O88168;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Orfde14.
OS Enterococcus faecalis (Streptococcus faecalis).

```

OC Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
 OX NCBI_TaxID=1351;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=OG1RF;
 RX MEDLINE=98380380; PubMed=9712783;
 RA Xu Y., Murray B.E., Weinstock G.M.;
 RT "A cluster of genes involved in polysaccharide biosynthesis from
 Enterococcus faecalis OG1RF".
 RL Infect. Immun. 66:4313-4323 (1998).
 DR EMBL; AF071085; AAC35928.1; -;
 DR GO; GO:0008757; F:S-adenosylmethionine-dependent methyltransf. . . ; IEA.
 DR InterPro; IPR001601; Methyltransf.
 DR InterPro; IPR000051; SAM bind.
 SQ SEQUENCE 392 AA; 44995 MW; 687A988FC2078CF6 CRC64;

Query Match 52.0%; Score 51; DB 2; Length 392;
 Best Local Similarity 47.1%; Pred. No. 22;
 Matches 8; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

Qy 1 DDPDVKEPDLQPFQVQ 17
 ||| :|:|:|:|:|:|:
 Db 256 DQVDLQKPTKQFOLK 272

RESULT 6
 Q832P0 PRELIMINARY; PRT; 1047 AA.
 AC Q832P0;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Glycosyl transferase, group 2 family protein.
 GN OrderedLocustNames=EF2181;
 OS Enterococcus faecalis (Streptococcus faecalis).
 OC Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
 OX NCBI_TaxID=1351;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=V583 / ATCC 700802;
 RX MEDLINE=22550857; PubMed=12663927; DOI=10.1126/science.1080613;
 RA Paulsen I.T., Banerjee L., Myers G.S.A., Nelson K.E., Seshadri R.,
 Read T.D., Fouts D.E., Eisen J.A., Gill S.R., Heidelberg J.F.,
 Tettelin H., Dodson R.J., Umayam L.A., Brinkac L.M., Beanan M.J.,
 Daugherty S.C., DeBoy R.T., Durkin S.A., Kolonay J.F., Madupu R.,
 Nelson W.C., Vamathevan J.J., Tran B., Upton J., Hansen T., Shetty J.,
 Khouri H.M., Utterback T.R., Radune D., Ketchum K.A., Dougherty B.A.,
 Fraser C.M.;
 RT "Role of mobile DNA in the evolution of vancomycin-resistant
 Enterococcus faecalis".
 RL Science 299:2071-2074 (2003).
 DR EMBL; AE016953; AA081913.1; -;
 DR TIGR; EF2181; -;
 DR GO; GO:0008757; F:S-adenosylmethionine-dependent methyltransf. . . ; IEA.
 DR InterPro; IPR001173; Glyco trans 2.
 DR InterPro; IPR001601; Methyltransf.
 DR InterPro; IPR000051; SAM bind.
 DR Pfam; PF00535; Glycosyltransf. 2; 1.
 KW Complete proteome; Transferase.
 SQ SEQUENCE 1047 AA; 119728 MW; 621F8B792F814E36 CRC64;

Query Match 52.0%; Score 51; DB 2; Length 1047;
 Best Local Similarity 47.1%; Pred. No. 63;
 Matches 8; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

Qy 1 DDPDVKEPDLQPFQVQ 17
 ||| :|:|:|:|:|:|:
 Db 256 DQVDLQKPTKQFOLK 272

RESULT 7

TTL_HUMAN

ID TTL_HUMAN STANDARD; PRT; 377 AA.
 AC Q8NG68; O72302; O8N426;
 DT 05-JUL-2004 (Rel. 44, Created)
 DT 05-JUL-2004 (Rel. 44, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Tubulin-tyrosine ligase (EC 6.3.2.25) (TTL).
 GN Name=TTL;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Miyazaki K., Okamoto Y., Kato C., Sakamoto M., Ohira M., Morohashi A.,
 Nakagawara A.;
 RT "Homo sapiens tubulin tyrosine ligase mRNA, complete cds.";
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
 Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 Butlerfield Y.S.N., Krzyzinski M.I., Skalska U., Smalls D.E.,
 Schnerch A., Schein J.E., Jones S.J.M., Maria M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [3]
 RP SEQUENCE OF 205-377 FROM N.A.
 RC TISSUE=Uterus;
 RA Lauber J., Bahr A., Mewes H.-W., Weil B., Amid C., Osanger A.,
 Fobo G., Han M., Wiemann S.;
 RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Catalyzes the posttranslational addition of a tyrosine
 to the C-terminal end of dephosphorylated alpha-tubulin (By
 similarity).
 CC -!- CATALYTIC ACTIVITY: ATP + dephosphorylated alpha-tubulin + L-tyrosine
 = alpha-tubulin + ADP + phosphate. (By similarity).
 CC -!- COFACTOR: Magnesium and potassium (By similarity).
 CC -!- SUBUNIT: Monomer (By similarity).
 CC -!- SIMILARITY: Contains 1 TTL domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AB071393; BAC06832.2; -;
 CC EMBL; BC036819; AAH36819.1; -;
 CC EMBL; BX538316; CAD98091.1; -;
 CC Genew: HGNC:21586; TTL.
 CC MIM; 608291; -;
 CC InterPro; IPR004344; Tub_tyr_lygase.
 CC Ligase; Magnesium; Potassium.
 FT DOMAIN 81 367 TTL.
 FT CONFLICT 245 245 I -> V (in Ref. 3).

```

Db      424  DFPPGVGEEDVGLPV 439
      ||| ||| ||| ||| |||
RESULT 9
Q835W9          PRELIMINARY;          PRT;          541 AA.
AC      Q835M9;
DT      01-JUN-2003 (TrEMBLrel. 24, Created)
DT      01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT      01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE      Glucan 1,6-alpha-glucosidase, putative.
GN      OrderedLocusNames=Brl348;
OS      Enterococcus faecalis (Streptococcus faecalis).
OC      Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus
OX      NCBI_TaxID=1351;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=V583 / ATCC 700802;
RX      MEDLINE=22550857; PubMed=12663927; DOI=10.1126/science.1080613;
RA      Paulsen I.T., Banerjee L., Myers G.S.A., Nelson K.E., Seshadri R.,
RA      Read T.D., Fouts D.E., Eisen J.A., Gill S.R., Heidelberg J.F.,
RA      Tettelin H., Dodson R.J., Unayam L.A., Brinkac L.M., Beanan M.J.,
RA      Daugherty S.C., DeBoy R.T., Durkin S.A., Kolonay J.F., Madupu R.,
RA      Nelson W.C., Vamathevan J.J., Tran B., Upton J., Hansen T., Shetty J.,
RA      Khouri H.M., Uutterback T.R., Radune D., Ketchum K.A., Dougherty B.A.,
RA      Fraser C.M.;
RT      "Role of mobile DNA in the evolution of vancomycin-resistant
RT      Enterococcus faecalis.";
RL      Science 299:2071-2074 (2003).
DR      EMBL; AE016951; AA081139.1; -.
DR      HSSP; P21332; 1UOK.
DR      TIGR; EF1348; -.
DR      GO; GO:0004556; P:alpha-amylase activity; IEA.
DR      GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR      InterPro; IPR006047; Alpha amyl cat.
DR      InterPro; IPR006589; Alp amyl_cat_sub.
DR      Pfam; PF00128; Alpha-amylase; 1.
DR      SMART; SM00642; Amy; 1.
DR      Complete proteome.
KW      SEQUENCE 541 AA; 62718 MW; ED0DB68653A7DC98 CRC64;
SQ
Query Match          49.0%; Score 48; DB 2; Length 541;
Best Local Similarity 62.5%; Pred. No. 89;
Matches 10; Conservative 1; Mismatches 5; Indels 0; Gaps
OY      1  DQPPDVEXPDLOPFQV 15
      ||| ||| ||| ||| |||
Db      274  DQPGKKEWDLOPMEV 289
      ||| ||| ||| ||| |||
RESULT 10
Q6CHL9          PRELIMINARY;          PRT;          652 AA.
AC      Q6CHL9;
DT      25-OCT-2004 (TrEMBLrel. 28, Created)
DT      25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT      25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE      Similarity.
GN      ORFNames=YALI0A07557g;
OS      Yarrowia lipolytica CL1899.
OC      Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OX      Saccharomycetales; Dipodascaceae; Yarrowia.
OX      NCBI_TaxID=284591;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=CL1899;
RG      Genolevures;
RA      Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
RA      Lafontaine I., de Montigny J., Marck C., Neuveglise C., Tallia E.,
RA      Goffard N., Franguel L., Aigle M., Anthouard V., Babour A., Barbe V.,
RA      Barnay S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,
RA      Boirame A., Bover J., Cattolico L., Confanioli F., de Daruvar A.,

```

```

RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
RA Hantrave F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA Nicaud J.M., Nikolski M., Oztas S., Ozier-Kalogeropoulos O.,
RA Pellenz S., Tekai F., Wesolowski-Louvel M., Suleau A.,
RA Swennen D., Tekaia F., Wesolowski-Louvel M., Westhof E., Wirth B.,
RA Zeniou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,
RA Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,
RA Winkler P., Souciet J.L.;
RA "Genome evolution in yeasts.";
RL Nature 430:35-44 (2004).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CLIB99;
RA Genoscope;
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR382127; CAG83768.1; -.
SQ SEQUENCE 652 AA; 74732 MW; 7C20193BD6CDE435 CRC64;

Query Match 49.0%; Score 48; DB 2; Length 652;
Best Local Similarity 53.8%; Pred. No. 1.1e+02;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 1 DQPDVEXKPDLP 13
Db 169 DGPPELRPELDP 181

RESULT 11
Q700G4
ID Q700G4 PRELIMINARY; PRT; 320 AA.
AC Q700G4;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Homeodomain transcription factor BW6C (Fragment).
GN Name=BW6C;
OS Ustilago maydis (Smut fungus).
OC Eukaryota; Fungi; Basidiomycota; Ustilaginomycetes;
OC Ustilaginomycetidae; Ustilaginales; Ustilaginaceae; Ustilago.
OX NCBI_TaxID=5270;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=All1;
RA Kaemper J.T., Friedrich M.W., Kahmann R.;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
DR EMBL; AJ630080; CAF34016.1; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001356; Homeobox_like.
DR Pfam; PF00046; Homeobox; 1.
DR PRINTS; PR00024; HOMEBOX.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00071; HOMEBOX 2; 1.
DR DNA-binding; Homeobox; Nuclear protein.
KW NON_TER 320 320
SQ SEQUENCE 320 AA; 36141 MW; A7BCC7EB713B107F CRC64;

Query Match 48.0%; Score 47; DB 2; Length 320;
Best Local Similarity 47.1%; Pred. No. 71;
Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 2 QPDVEXKPDLPQFQVQS 18
Db 217 EPTDSTQPDLSPPRSES 233

RESULT 12
Q700G6
ID Q700G6 PRELIMINARY; PRT; 320 AA.
AC Q700G6;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Homeodomain transcription factor BW6b (Fragment).
GN Name=BW6b;
OS Ustilago maydis (Smut fungus).
OC Eukaryota; Fungi; Basidiomycota; Ustilaginomycetes;
OC Ustilaginomycetidae; Ustilaginales; Ustilaginaceae; Ustilago.
OX NCBI_TaxID=5270;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Umb1;
RA Kaemper J.T., Friedrich M.W., Kahmann R.;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
DR EMBL; AJ630078; CAF34012.1; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001356; Homeobox_like.
DR Pfam; PF00046; Homeobox; 1.
DR PRINTS; PR00024; HOMEBOX.
DR SMART; SM00389; HOX; 1.
DR PRODOM; PD000010; Homeobox; 1.
DR PROSITE; PS00071; HOMEBOX 2; 1.
DR DNA-binding; Homeobox; Nuclear protein.
KW NON_TER 320 320
SQ SEQUENCE 320 AA; 36141 MW; A7BCC7EB713B107F CRC64;

Query Match 48.0%; Score 47; DB 2; Length 320;
Best Local Similarity 47.1%; Pred. No. 71;
Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 2 QPDVEXKPDLPQFQVQS 18
Db 217 EPTDSTQPDLSPPRSES 233

RESULT 13
Q700G8
ID Q700G8 PRELIMINARY; PRT; 320 AA.
AC Q700G8;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Homeodomain transcription factor BW6a (Fragment).
GN Name=BW6a;
OS Ustilago maydis (Smut fungus).
OC Eukaryota; Fungi; Basidiomycota; Ustilaginomycetes;
OC Ustilaginomycetidae; Ustilaginales; Ustilaginaceae; Ustilago.
OX NCBI_TaxID=5270;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Umb1;
RA Kaemper J.T., Friedrich M.W., Kahmann R.;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
DR EMBL; AJ630078; CAF34012.1; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001356; Homeobox_like.
DR Pfam; PF00046; Homeobox; 1.
DR PRINTS; PR00024; HOMEBOX.
DR SMART; SM00389; HOX; 1.
DR PRODOM; PD000010; Homeobox; 1.
DR PROSITE; PS00071; HOMEBOX 2; 1.
DR DNA-binding; Homeobox; Nuclear protein.
KW NON_TER 320 320
SQ SEQUENCE 320 AA; 36137 MW; 227986FF213D4E06 CRC64;

Query Match 48.0%; Score 47; DB 2; Length 320;
Best Local Similarity 47.1%; Pred. No. 71;
Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 2 QPDVEXKPDLPQFQVQS 18
Db 217 EPTDSTQPDLSPPRSES 233

```

```
SQ SEQUENCE 320 AA; 36049 MW; A0EFF383413E735A CRC64;
Query Match 48.0%; Score 47; DB 2; Length 320;
Best Local Similarity 47.1%; Pred. No. 71;
Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 2 QPDPVEKPDLPQFQVQS 18
Db 217 EPTDSTQPDLSPPRSSES 233

RESULT 14
Q70019 PRELIMINARY; PRT; 320 AA.
AC Q70019;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Homeodomain transcription factor Bw11 (Fragment).
GN Name=Bw11;
OS Ustilago maydis (Smut fungus).
OC Eukaryota; Fungi; Basidiomycota; Ustilaginomycetes;
OC Ustilaginomycetidae; Ustilaginales; Ustilaginaceae; Ustilago.
OX NCBI_TaxID=5270;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC22896;
RA Kaemper J.T., Friedrich M.W., Kahmann R.;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
DR EMBL; AJ6330067; CAF33990.1; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR009057; Homeobox.
DR PRINTS; PR00024; HOMEBOX.
DR Pfam; PF00046; Homeobox; 1.
DR ProDom; PD00010; Homeobox.
DR PRINTS; PR00024; HOMEBOX.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS50071; HOMEBOX 2; 1.
KW DNA-binding; Homeobox; Nuclear protein.
FT NON TER 320 320
SQ SEQUENCE 320 AA; 35989 MW; 25061B961594DAC2 CRC64;

Query Match 48.0%; Score 47; DB 2; Length 320;
Best Local Similarity 47.1%; Pred. No. 71;
Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 2 QPDPVEKPDLPQFQVQS 18
Db 217 EPTDSTQPDLSPPRSSES 233

RESULT 15
Q99118 PRELIMINARY; PRT; 404 AA.
AC Q99118;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Bw4 protein (Fragment).
GN Name=Bw4;
OS Ustilago maydis (Smut fungus).
OC Eukaryota; Fungi; Basidiomycota; Ustilaginomycetes;
OC Ustilaginomycetidae; Ustilaginales; Ustilaginaceae; Ustilago.
OX NCBI_TaxID=5270;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RK138;
RA Gillissen B., Bergemann J., Sandmann C., Schroeer B., Boelker M.,
MEDLINE=92154679; PubMed=1739973; DOI=10.1016/0092-8674(92)90141-X;
```

```
RT "A two-component regulatory system for self/non-self recognition in
RL Ustilago maydis."
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
DR EMBL; M84181; AAA34223.1; -.
DR PIR; D42094; D42094.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001356; Homeobox.
DR InterPro; IPR009057; Homeodomain_like.
DR Pfam; PF00046; Homeobox; 1.
DR PRINTS; PR00024; HOMEBOX.
DR ProDom; PD00010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS50071; HOMEBOX 2; 1.
KW DNA-binding; Homeobox; Nuclear protein.
FT NON TER 1 1
SQ SEQUENCE 404 AA; 45439 MW; 4B2C71857AC82910 CRC64;

Query Match 48.0%; Score 47; DB 2; Length 404;
Best Local Similarity 47.1%; Pred. No. 92;
Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 2 QPDPVEKPDLPQFQVQS 18
Db 198 EPTDSTQPDLSPPRSSES 214

Search completed: April 12, 2005, 08:05:39
Job time : 87.4182 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 12, 2005, 07:49:31 ; Search time 67.2 Seconds
(without alignments)
69.064 Million cell updates/sec

Title: US-10-691-330-4
Perfect score: 62
Sequence: 1 LFFPLPVNVLP 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	62	100.0	12	4	AAB72503 Colostrin
2	62	100.0	12	4	AAB59323 Ewe colos
3	62	100.0	12	4	AAB72249 Colostrin
4	62	100.0	12	4	AAB72535 Colostrin
5	62	100.0	12	5	AAO14580 Neural ce
6	62	100.0	12	5	AAO14580 Colostrin
7	62	100.0	12	5	AAO14580 Colostrin
8	62	100.0	12	8	ADN60298 Constitue
9	62	100.0	12	8	ADN60298 Constitue
10	62	100.0	12	8	ADN60298 Constitue
11	42	67.7	166	3	AAB59353 Ewe colos
12	42	67.7	179	3	AAG33579 Arabidops
13	42	67.7	221	3	AAG33578 Arabidops
14	41	66.1	104	7	AAG33577 Arabidops
15	41	66.1	126	6	ADN6263 E. faeciu
16	41	66.1	692	6	ADN6263 E. faeciu
17	41	66.1	692	6	ADN6263 E. faeciu
18	40	64.5	733	6	ADN6263 E. faeciu
19	39	62.9	243	3	ADN6263 E. faeciu
20	39	62.9	422	6	ADN6263 E. faeciu
21	39	62.9	528	7	ADN6263 E. faeciu
22	39	62.9	710	5	ADN6263 E. faeciu
23	39	62.9	750	3	ADN6263 E. faeciu
24	38	61.3	113	6	ADN6263 E. faeciu
25	38	61.3	122	6	ADN6263 E. faeciu

ALIGNMENTS

RESULT 1

AAB72503 ID AAB72503 standard; peptide; 12 AA.

XX AC AAB72503;

DT 09-MAY-2001 (first entry)

DE Colostrin peptide #4.

XX KW Dermatological; oxidative stress regulator; colostrin.

XX OS Unidentified.

XX PN WO200112650-A2.

XX PD 22-FEB-2001.

XX PF 17-AUG-2000; 2000WO-US022665.

XX PR 17-AUG-1999; 99US-0149310P.

XX PA (TEXA) UNIV TEXAS SYSTEM.

XX PI Stanton GJ, Hughes TK, Boldogh I;

XX DR WPI; 2001-218342/22.

XX PT Modulating oxidative stress level in a cell, involves contacting the cell with an oxidative stress regulator selected from colostrin, its constituent peptide, analog or their combinations.

XX PS Claim 6; Page 25; 48pp; English.

XX CC The present invention relates to a method for modulating the oxidative stress level in a cell or a patient, comprising contacting the cell with, or administering to the patient, an oxidative stress regulator selected from colostrin, or its constituent peptide (e.g. the present peptide), to change the level of an oxidising species in the cell. The method can be used to treat oxidative damage to skin, by decreasing or preventing an increase in the level of damage to a biomolecule of the patient

XX SQ Sequence 12 AA;

Query Match 100.0%; Score 62; DB 4; Length 12;

Best Local Similarity 100.0%; Pred. No. 0.00034;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LFFFLPVNVLP 12
 DB 1 LFFFLPVNVLP 12

RESULT 2

AAB59323
 ID AAB59323 standard; peptide; 12 AA.

XX AC AAB59323;

XX DT 21-MAR-2001 (first entry)

XX DE Ewe colostrinin peptide fragment B-8.

XX KW Sheep; colostrinin; proline rich polypeptide; colostrum; immune disorder;
 XX central nervous system disorder; dietary supplement; beta-amyloid plaque.

XX OS Ovis sp.

XX PN WO200075173-A2.

XX PD 14-DEC-2000.

XX PF 02-JUN-2000; 2000WO-GB002128.

XX PR 02-JUN-1999; 95GB-00012852.

XX PA (REGE-) REGEN THERAPEUTICS PLC.

XX PI Georgiades JA;

XX DR WPI; 2001-071058/08.

XX PT Peptides having an N-terminal amino acid sequence isolated from
 PT colostrinin for treating e.g. disorders of the central nervous system and
 PT immune system, viral and bacterial infections, and diseases characterized
 PT by amyloid plaques.

PS Claim 7; Page 27; 63pp; English.

XX CC The present invention provides the sequences of a number of peptides
 CC found in ewe's colostrum. Colostrinin is the proline-rich polypeptide
 CC fragment of colostrum. These peptides can be used in the treatment of
 CC central nervous system disorders such as senile dementia, Parkinson's
 CC disease, Alzheimer's disease, psychosis and neurosis, immune system
 CC disorders such as bacterial and viral infections, to improve the
 CC development of a child's immune system, as a dietary supplement, and to
 CC promote the dissolution of beta-amyloid plaques

XX SQ Sequence 12 AA;

Query Match 100.0%; Score 62; DB 4; Length 12;
 Best Local Similarity 100.0%; Pred. No. 0.00054;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LFFFLPVNVLP 12
 DB 1 LFFFLPVNVLP 12

RESULT 3

AAB72249
 ID AAB72249 standard; peptide; 12 AA.

XX AC AAB72249;

XX DT 14-MAY-2001 (first entry)

XX DE Colostrinin derived cytokine inducing peptide SEQ ID 4.

XX KW Colostrinin; immune response; cytokine; blood cell proliferation;
 KW central nervous system disorder; neurological disorder; mental disorder;

KW dementia; neurodegenerative disease; Alzheimer's disease; psychosis;
 KW neurosis; infection.

XX OS Synthetic.

XX PN WO200111937-A2.

XX PD 22-FEB-2001.

XX PF 17-AUG-2000; 2000WO-US022818.

XX PR 17-AUG-1999; 99US-0149311P.

XX PA (TEXA) UNIV TEXAS SYSTEM.

XX PA (REGE-) REGEN THERAPEUTICS PLC.

XX PI Stanton GJ, Hughes TK, Boldogh I, Georgiades J;

XX DR WPI; 2001-202804/20.

XX PT Inducing a cytokine and modulating an immune response, useful for
 PT treating central nervous system diseases and bacterial and viral
 PT infections, comprises administering colostrinin as an immunological
 PT regulator.

PS Claim 1; Page 34; 50pp; English.

XX CC Sequences AAB72246 - AAB72275 represent peptides derived from colostrinin,
 CC a proline rich polypeptide aggregate contained in colostrum. The peptides
 CC have immune response modulatory activity, and are capable of inducing
 CC cytokines. Colostrinin and its derived peptides are useful for inducing
 CC cytokine production, for modulating an immunological response and for
 CC inducing blood cell proliferation. The peptides are useful in the
 CC treatment of disorders of the central nervous system, neurological
 CC disorders, mental disorders, dementia, neurodegenerative diseases,
 CC Alzheimer's disease, motor neurone disease, psychosis, neurosis, chronic
 CC disorders of the immune system, bacterial and viral infections and
 CC acquired immunological deficiencies

XX SQ Sequence 12 AA;

Query Match 100.0%; Score 62; DB 4; Length 12;
 Best Local Similarity 100.0%; Pred. No. 0.00054;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LFFFLPVNVLP 12
 DB 1 LFFFLPVNVLP 12

RESULT 4

AAB72535
 ID AAB72535 standard; peptide; 12 AA.

XX AC AAB72535;

XX DT 09-MAY-2001 (first entry)

XX DE Colostrinin peptide #4.

XX KW Neuroprotective; neural cell differentiation regulator; colostrinin;
 KW colostrum.

XX OS Unidentified.

XX PN WO200112651-A2.

XX PD 22-FEB-2001.

XX PF 17-AUG-2000; 2000WO-US022774.

XX PR 17-AUG-1999; 99US-0149633P.

XX XX

PA (TEXA) UNIV TEXAS SYSTEM.
 PI Boldogh I;
 XX WPI; 2001-226545/23.
 DR
 XX Use of colostrinin, its constituent peptide or analog as a neural cell
 PT regulator, for promoting neural cell differentiation and treating damaged
 PT neural cells in a patient.
 XX
 XX Claim 6; Page 21; 35pp; English.
 PS
 XX The present invention relates to a method for promoting neural cell
 CC differentiation and treating damaged neural cells, using colostrinin and
 CC colostrinin constituent peptides (e.g. the present peptide) as a neural
 CC cell regulator. Colostrinin is a polypeptide complex found in colostrum
 XX
 XX Sequence 12 AA;
 SQ

Query Match 100.0%; Score 62; DB 4; Length 12;
 Best Local Similarity 100.0%; Pred. No. 0.00054;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LFFFLPVNVLP 12
 Db 1 LFFFLPVNVLP 12
 |||||

RESULT 5
 AA014580
 ID AA014580 standard; peptide; 12 AA.
 XX
 AC AA014580;
 XX
 DT 27-MAY-2002 (first entry)
 XX
 DE Neural cell regulatory colostrinin peptide 4.
 XX
 KW Neural cell differentiation; neural cell regulator; colostrinin peptide;
 KW neural cell formation; proline-rich polypeptide aggregate; colostrum;
 KW neural cell treatment.
 XX
 OS Unidentified.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 12 /note= "Optional C-terminal amide"
 FT
 FN WO200213851-A1.
 XX
 PD 21-FEB-2002.
 XX
 PF 17-AUG-2000; 2000WO-US022777.
 XX
 PR 17-AUG-2000; 2000WO-US022777.
 XX
 XX (TEXA) UNIV TEXAS SYSTEM.
 XX Boldogh I, Stanton JG, Hughes TK;
 XX WPI; 2002-269152/31.
 XX Promoting cell differentiation in a patient involves use of blood cell
 PT regulator selected from colostrinin, its constituent peptide and/or
 PT analog.
 XX
 XX Claim 7; Page 21; 37pp; English.
 PS
 XX The invention comprises a method for promoting cell differentiation (e.g.
 CC neural cell differentiation). The method involves contacting cells with a
 CC neural cell regulator (i.e. a colostrinin peptide) in order to change the
 CC cells in morphology to form neural cells. Colostrinin is a proline-rich
 CC polypeptide aggregate that is present in colostrum. The method of the

CC invention is useful for promoting the differentiation of cells and for
 CC treating damaged neural cells in a patient. The present amino acid
 CC sequence represents a specifically claimed colostrinin peptide used in
 CC the method of the invention
 XX
 SQ Sequence 12 AA;
 Query Match 100.0%; Score 62; DB 5; Length 12;
 Best Local Similarity 100.0%; Pred. No. 0.00054;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LFFFLPVNVLP 12
 Db 1 LFFFLPVNVLP 12
 |||||

RESULT 6
 AA014580
 ID AA014580 standard; peptide; 12 AA.
 XX
 AC AA014580;
 XX
 DT 30-MAY-2002 (first entry)
 XX
 DE Colostrinin constituent peptide.
 XX
 KW Colostrinin; colostrum; immunomodulator; cardiovascular;
 KW blood cell regulator; cytokine inducer; beta-casein; human.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 12 /note= "Optional C-terminal amidation"
 FT
 FN WO200213849-A1.
 XX
 PD 21-FEB-2002.
 XX
 PF 17-AUG-2000; 2000WO-US022775.
 XX
 PR 17-AUG-2000; 2000WO-US022775.
 XX
 XX (TEXA) UNIV TEXAS SYSTEM.
 XX (REGG-) REGEN THERAPEUTICS PLC.
 XX Stanton GJ, Hughes TK, Boldogh I, Georgiades J;
 XX WPI; 2002-269150/31.
 XX Modulation of blood cell proliferation in a patient involves use of blood
 PT cell regulator selected from colostrinin, its constituent peptide and/or
 PT analog.
 XX
 XX Claim 1; Page 34; 54pp; English.
 PS
 XX The present sequence is that of a colostrinin constituent peptide that is
 CC preferred for use as an immunological regulator and as a blood cell
 CC regulator in claimed methods of the invention. It is classified as having
 CC a beta-casein homologue precursor. Methods are claimed for: inducing a
 CC cytokine in a cell by contact with an immunological regulator, where the
 CC cell is present in a cell culture, a tissue, an organ or an organism, and
 CC the cell is mammalian, including human; modulating an immune response in
 CC a cell by contact with the immunological regulator under conditions
 CC effective to induce a cytokine; modulating an immune response in a
 CC patient by administering an immunological regulator under conditions
 CC effective to induce a cytokine, where the immunological regulator is
 CC administered topically or as part of a dietary supplement, and where the
 CC immune response is specific or non specific, an interferon response or an
 CC antibody response; modulating blood cell proliferation by contacting
 CC blood cells with a blood cell regulator, where the blood cells are
 CC present in a cell culture or an organism, are mammalian or human, and a
 CC where the blood cells are increased in number or differentiated; and a

CC method for modulating blood cell proliferation in a patent. A claimed
 CC cytokine-inducing composition comprises a pharmaceutical carrier and an
 CC active agent such as the present peptide. Cytokines induced by this
 CC peptide in human leucocyte cultures include interferon-gamma, tumour
 CC necrosis factor-alpha, interleukin-6 and interleukin-10
 XX

SQ Sequence 12 AA;
 Query Match 100.0%; Score 62; DB 5; Length 12;
 Best Local Similarity 100.0%; Pred. No. 0.00054;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 LFFFLPVNVLP 12
 |||||
 Db 1 LFFFLPVNVLP 12
 |||||

RESULT 7
 AAE20231
 ID AAE20231 standard; peptide; 12 AA.
 XX
 AC AAE20231;
 XX
 DT 18-JUN-2002 (first entry)
 XX
 DE Colostrinin constituent peptide #4.
 XX
 KW Blood cell regulator; colostrinin; constituent peptide; oxidative stress;
 KW therapy; oxidative damage; skin; aging; wound healing; cell replacement;
 KW tissue; organ; cosmetic procedure; repair; regeneration; preservation;
 KW transplantation; implantation; dermatological; vulnery.
 XX
 OS Unidentified.

XX Key Location/Qualifiers
 FT Modified-site 12 /note= "Optionally C-terminal amide"
 FT
 FN WO200213850-A1.
 XX
 XX 21-FEB-2002.
 XX
 XX 17-AUG-2000; 2000WO-US022776.
 XX
 XX 17-AUG-2000; 2000WO-US022776.
 XX
 XX (TEXA) UNIV TEXAS SYSTEM.
 XX
 XX Stanton GU, Hughes TK, Boldogh I;
 XX WPI; 2002-269151/31.
 XX
 XX Composition useful for the modulation of blood cell proliferation in a
 XX patient comprises a blood cell regulator selected from colostrinin, its
 XX constituent peptide and/or analog.
 XX
 XX Claim 6; Page 25; 51pp; English.

XX The invention relates to a composition which comprises a blood cell
 XX regulator selected from colostrinin, its constituent peptide and/or
 XX analogue. The invention is used for modulating the oxidative stress level
 XX in a cell e.g. mammalian or human cell present in a cell culture, tissue,
 XX organ, or organism; or for treating oxidative damage to the skin of a
 XX patient e.g. animal or human; to modulate oxidative stress during/ after
 XX a premature birth or normal birth, preventing/delaying aging in a
 XX patient, enhancing wound healing, and the reduction of side effects of
 XX cosmetic procedures. The method changes the level of an oxidising species
 XX in the cell, such as decreases or prevents increase in the level of
 XX damage to a biomolecule of the patient selected from DNA, protein and/or
 XX lipid, compared to the same conditions when the oxidative stress
 XX regulator is not present. The modulation of oxidative stress results in
 XX enhanced repair, regeneration, and replacement of cells, tissues and
 XX organs (e.g. kidney, liver, pancreas, skin, and the other internal and

CC external organs), as well as enhanced preservation of such organs for
 CC transplantation, implantation, or scientific research. The present
 CC sequence is a colostrinin constituent peptide
 XX

SQ Sequence 12 AA;
 Query Match 100.0%; Score 62; DB 5; Length 12;
 Best Local Similarity 100.0%; Pred. No. 0.00054;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 LFFFLPVNVLP 12
 |||||
 Db 1 LFFFLPVNVLP 12
 |||||

RESULT 8
 ADN60298
 ID ADN60298 standard; peptide; 12 AA.
 XX
 AC ADN60298;
 XX
 DT 29-JUL-2004 (first entry)
 XX
 DE Constituent peptide of colostrinin SEQ ID NO:4.
 XX
 KW modulator; colostrinin; intracellular signaling molecule modulator;
 KW 4HNE-mediated lipid peroxidation down regulation; apoptosis inhibition;
 KW DNA damage; beta-amyloid; retinoic acid; cytostatic; 4HNE inhibitor;
 KW 4HNE-protein adduct formation reduction;
 KW 4HNE-mediated glutathione depletion inhibition; p53 inhibition;
 KW c-Jun NH2-terminal kinase inhibition.

XX Synthetic.
 XX
 XX WO2004037851-A2.
 XX
 XX 06-MAY-2004.
 XX
 XX 22-OCT-2003; 2003WO-US033423.
 XX
 XX 22-OCT-2002; 2002US-0420369P.
 XX
 XX (TEXA) UNIV TEXAS SYSTEM.
 XX (BOLD/) BOLDOGH I.
 XX (STAN/) STANTON J G.
 XX (GEOR/) GEORGIADIS J A.
 XX (HUGH/) HUGHES T K.
 XX (KRUZ/) KRUZEL M.

XX Boldogh I, Stanton JG, Georgiades JA, Hughes TK, Kruzel M;
 XX WPI; 2004-365494/34.
 XX
 XX Use of colostrinin for e.g. modulating an intracellular signaling
 XX molecule in a cell, down regulating 4HNE-mediated lipid peroxidation in a
 XX cell, inhibiting apoptosis in a cell, or protecting against DNA damage in
 XX a cell.

XX Claim 6; SEQ ID NO 4; 46pp; English.
 XX
 XX The present invention describes the use of a modulator selected from
 XX colostrinin, its constituent peptide, its active analogue, and a
 XX combination of these, for modulating an intracellular signaling molecule
 XX in a cell, down regulating 4HNE-mediated lipid peroxidation in a cell,
 XX inhibiting apoptosis in a cell, protecting against DNA damage in a cell,
 XX and reducing the toxic effect of beta-amyloid or retinoic acid on a cell.
 XX The modulator has cytostatic activity, and can be used as a 4HNE
 XX inhibitor. The modulator is useful in the manufacture of a medicament for
 XX reducing 4HNE-protein adduct formation, inhibiting 4HNE-mediated
 XX glutathione depletion, inhibiting 4HNE-induced activation of p53 protein,
 XX and/or inhibiting 4HNE-induced activation of c-Jun NH2-terminal kinases.
 XX Colostrinin, or its constituent peptide or active analogue is useful for
 XX inhibiting apoptosis in a cell, protecting against DNA damage in a cell,

CC or reducing the toxic effect of beta-amyloid or retinoic acid on a cell.
CC The present sequence represents a synthetic constituent peptide of
CC colostrinin, which can be used as a modulator in the present invention.

XX SQ Sequence 12 AA;
Query Match 100.0%; Score 62; DB 8; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.00054;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LFFFLPVVNVLP 12
Db 1 LFFFLPVVNVLP 12

RESULT 9
ADS74399
ID ADS74399 standard; peptide; 12 AA.

XX AC ADS74399;
XX XX
DT 16-DEC-2004 (first entry)
XX XX
DE Ovine colostrinin peptide.
XX XX
KW Colostrum; colostrinin; sheep; peptide purification.
XX XX
OS Ovis aries.

XX PN WO2004081038-A1.

XX PD 23-SEP-2004.

XX PF 10-MAR-2004; 2004WO-GB001014.

XX PR 11-MAR-2003; 2003GB-00005552.

XX PR 08-MAR-2004; 2004GB-00005190.

XX PA (REGG-) REGEN THERAPEUTICS PLC.

XX PI Georgiades JA, Polanowski A, Wilusz T, Kruzel ML;

XX DR WPI; 2004-677519/66.

XX PT Recovering peptides such as colostrinin from mammalian colostrum, by
PT mixing colostrum with alcohol to form alcohol phase containing peptides
PT and precipitate, separating alcohol phase from precipitate, and
PT recovering alcohol phase.

XX PS Example; SEQ ID NO 16; 41pp; English.

XX CC The present sequence is that of a peptide that can be recovered from
XX CC ovine colostrinin using the method of the invention. The invention
XX CC provides a method for the recovery of peptides (especially colostrinin)
XX CC from colostrum in substantially pure, biologically active form and in
XX CC high yield. The method involves mixing the colostrum with an alcohol to
XX CC form an alcohol phase containing the colostrinin and a precipitate
XX CC containing higher molecular weight caseins and other proteins. Best
XX CC results are obtained using methanol or ethanol of at least 80%, and
XX CC preferably up to 100%, purity. The alcohol phase is then separated from
XX CC the precipitate, and the colostrinin is separated from the alcohol,
XX CC preferably by evaporation, to form a colostrinin-rich phase, which is
XX CC recovered. A precipitation agent, such as ammonium sulfate, may be added
XX CC either to the alcohol phase or, preferably, to the colostrinin-rich phase
XX CC to induce precipitation of the colostrinin peptides. The method is
XX CC generally applicable to the separation of peptides from fluids containing
XX CC higher molecular weight proteins, lipids, carbohydrates and/or nucleic
XX CC acids. In an example from the invention, the antigenic profile of
XX CC peptides recovered from sheep colostrum using the alcohol precipitation
XX CC methods was determined by ELISA using antibodies prepared against 9
XX CC synthetic peptides, including a peptide having the present sequence
XX CC (denoted antigen class B-9).

SQ Sequence 12 AA;

Query Match 100.0%; Score 62; DB 8; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.00054;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LFFFLPVVNVLP 12
Db 1 LFFFLPVVNVLP 12

RESULT 10
AAB59353
ID AAB59353 standard; peptide; 14 AA.

XX AC AAB59353;

XX DT 21-MAR-2001 (first entry)

XX DE Ewe colostrinin peptide fragment derived sequence #13.

XX KW Sheep; colostrinin; proline rich polypeptide; colostrum; immune disorder;
XX KW central nervous system disorder; dietary supplement; beta-amyloid plaque.

XX OS Ovis sp.

XX PN WO200075173-A2.

XX PD 14-DEC-2000.

XX PF 02-JUN-2000; 2000WO-GB002128.

XX PR 02-JUN-1999; 99GB-00012852.

XX PA (REGG-) REGEN THERAPEUTICS PLC.

XX PI Georgiades JA;

XX DR WPI; 2001-071058/08.

XX PT Peptides having an N-terminal amino acid sequence isolated from
XX PT colostrinin for treating e.g. disorders of the central nervous system and
XX PT immune system, viral and bacterial infections, and diseases characterized
XX PT by amyloid plaques.

XX PS Claim 8; Page 27; 63pp; English.

XX CC The present invention provides the sequences of a number of peptides
XX CC found in ewe's colostrinin. Colostrinin is the proline-rich polypeptide
XX CC fragment of colostrum. These peptides can be used in the treatment of
XX CC central nervous system disorders such as senile dementia, Parkinson's
XX CC disease, Alzheimer's disease, psychosis and neurosis, immune system
XX CC disorders such as bacterial and viral infections, to improve the
XX CC development of a child's immune system, as a dietary supplement, and to
XX CC promote the dissolution of beta-amyloid plaques

SQ Sequence 14 AA;

Query Match 100.0%; Score 62; DB 4; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.00063;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LFFFLPVVNVLP 12
Db 2 LFFFLPVVNVLP 13

RESULT 11
AAG33579
ID AAG33579 standard; protein; 166 AA.

XX AC AAG33579;

XX XX


```
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 67.7%; Score 42; DB 3; Length 166;
Best Local Similarity 70.0%; Pred. No. 25;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 FFFLPVNVVL 11
Db 135 YFFLPVIXNL 144

RESULT 12
AAG33578
ID AAG33578 standard; protein; 179 AA.
XX AC
XX AAG33578;
XX AC
DT 18-OCT-2000 (first entry)
XX Arabidopsis thaliana protein fragment SEQ ID NO: 40710.
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
```

```
XX Arabidopsis thaliana.
OS EPI033405-A2.
PN 06-SEP-2000.
XX 25-FEB-2000; 2000EP-00301439.
XX 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 18-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
```


PF 25-FEB-2000; 2000EP-00301439.
XX 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-01213180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-013448P.
PR 30-APR-1999; 99US-0134048P.
PR 30-APR-1999; 99US-0134207P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0136229P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144086P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 06-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.

```

PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159233P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 14-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 67.7%; Score 42; DB 3; Length 221;
Best Local Similarity 70.0%; Pred. No. 34;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 FFFLPVNVNVL 11
DB 190 YFFFLPVNXL 199
:|||||:|

RESULT 14
ADC96263
ID ADC96263 standard; protein; 104 AA.
XX AC
XX ADC96263;
XX XX
XX 01-JAN-2004 (first entry)
XX DE
XX E. faecium protein sequence SEQ ID 5890.
XX XX
XX Vaccine; urinary tract infection; bacteraemia; endocarditis; wound;
XX KW abdominal-pelvic infection.
XX KW
XX OS Enterococcus faecium.
XX PN
XX US6583275-B1.
XX PD
XX 24-JUN-2003.
XX XX
XX 30-JUN-1998; 98US-00107532.
XX PF
XX 02-JUL-1997; 97US-0051571P.
XX PR
XX 14-MAY-1998; 98US-0085598P.
XX XX
XX (GENO-) GENOME THERAPEUTICS CORP.
XX PA
XX Doucette-Stamm LA, Bush D;
XX PI

```

```

XX WPI: 2003-799836/75.
DR N-PSDB; ADC92609.
XX
PT New isolated nucleic acid derived from Enterococcus faecium encoding an
PT Enterococcus faecium polypeptide useful for detection, prevention and
PT treatment of a pathological condition resulting from a bacterial
PT infection.
XX
PS Example 1; SEQ ID NO 5890; 243pp; English.
XX
XX The invention relates to an isolated nucleic acid derived from
CC Enterococcus faecium encoding an Enterococcus faecium polypeptide having
CC one of 10 fully defined sequences given in the (or comprising 40
CC sequential nucleotides chosen from any of the nucleic acids, its
CC complement or sequences hybridising to it). Also included are a
CC recombinant vector comprising the nucleic acid operably linked to a
CC transcription regulatory element, a cell comprising the vector and a
CC single-stranded probe comprising the nucleic acid. The nucleic acids are
CC chosen from 3654 disclosed sequences encoding 3654 disclosed proteins.
CC The nucleic acids is useful for diagnosing pathological conditions
CC resulting from E. faecium bacterial infection (e.g. urinary tract
CC infection, bacteraemia, endocarditis, wounds and abdominal-pelvic
CC infection) and for screening drugs such as agonists and antagonists. The
CC nucleic acid is useful for recombinant production of Candida albicans -
CC derived peptides or antisense polypeptides. Pharmaceutical compositions
CC and vaccines containing the nucleic acid are useful for preventing or
CC treating Enterococcus faecium infections. The present sequence represents
CC one if the disclosed E. faecium proteins.
XX
SQ Sequence 104 AA;
Query Match 66.1%; Score 41; DB 7; Length 104;
Best Local Similarity 66.7%; Pred. No. 23;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 FFFLPVNVN 10
DB 91 FFFLIPLNV 99
:|||||:|

RESULT 15
ADA54377
ID ADA54377 standard; protein; 126 AA.
XX AC
XX ADA54377;
XX XX
XX 20-NOV-2003 (first entry)
XX DE
XX Human protein, SEQ ID 1945.
XX XX
XX Cytostatic; Anti-inflammatory; Osteopathic; Neuroprotective; Nootropic;
XX KW Gene Therapy; human; secretory protein; membrane proteins; cancer;
XX KW inflammatory disease; osteoporosis; neurological disease.
XX XX
XX Homo sapiens.
XX XX
XX EPI293569-A2.
XX PN
XX 19-MAR-2003.
XX PD
XX 21-MAR-2002; 2002EP-00006586.
XX PF
XX 14-SEP-2001; 2001JP-00328381.
XX PR
XX 24-JAN-2002; 2002US-0350435P.
XX PR
XX (HELI-) HELIX RES INST.
XX PA (REAS-) RES ASSOC BIOTECHNOLOGY.
XX FA
XX Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
XX PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
XX FI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;
XX XX

```

Search completed: April 12, 2005, 08:18:19
Job time : 69.2 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 12, 2005, 07:49:30 ; Search time 14.0727 Seconds
(without alignments)
63.654 Million cell updates/sec

Title: US-10-691-330-4
Perfect score: 62
Sequence: 1 LFFFLPVNVLP 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues
Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/1/iaa/5A-COMB.pap.*
2: /cgn2_6/ptodata/1/iaa/5B-COMB.pap.*
3: /cgn2_6/ptodata/1/iaa/6A-COMB.pap.*
4: /cgn2_6/ptodata/1/iaa/6B-COMB.pap.*
5: /cgn2_6/ptodata/1/iaa/PCTUS-COMB.pap.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	62	100.0	12	4	US-09-641-803-4
2	41	66.1	104	4	US-09-107-532A-5890
3	40	64.5	235	4	US-09-270-767-35554
4	40	64.5	235	4	US-09-270-767-50771
5	40	64.5	733	4	US-09-328-322-6242
6	39	62.9	528	4	US-09-252-991A-23368
7	38	61.3	483	4	US-09-489-039A-7429
8	37	59.7	42	4	US-09-205-258-456
9	37	59.7	72	4	US-09-107-532A-6843
10	37	59.7	103	4	US-09-205-258-251
11	37	59.7	405	4	US-09-489-039A-12853
12	36	58.1	48	4	US-09-071-035-196
13	36	58.1	109	4	US-09-071-035-194
14	36	58.1	123	4	US-09-134-000C-4445
15	36	58.1	137	4	US-09-489-039A-11180
16	36	58.1	174	4	US-09-311-021-138
17	36	58.1	307	2	US-08-782-760-6
18	36	58.1	307	5	PCT-US96-00995-6
19	36	58.1	348	4	US-09-902-540-16439
20	36	58.1	396	1	US-07-649-591B-4
21	36	58.1	396	1	US-08-277-540-4
22	36	58.1	396	1	US-08-430-787A-4
23	36	58.1	424	4	US-09-134-000C-5836
24	36	58.1	440	4	US-09-902-540-15770
25	36	58.1	455	4	US-09-248-796A-14873
26	36	58.1	585	4	US-09-248-796A-14925
27	36	58.1	918	4	US-09-438-185A-999

28	36	58.1	1604	4	US-09-949-016-11186	Sequence 11186, A
29	36	58.1	2555	4	US-09-538-092-820	Sequence 820, App
30	35	56.5	60	4	US-09-248-796A-23666	Sequence 23666, A
31	35	56.5	164	4	US-09-248-796A-16166	Sequence 16166, A
32	35	56.5	204	4	US-09-134-000C-3659	Sequence 3659, App
33	35	56.5	227	4	US-09-904-615-126	Sequence 126, App
34	35	56.5	240	4	US-09-270-767-37747	Sequence 37747, A
35	35	56.5	240	4	US-09-270-767-52964	Sequence 52964, A
36	35	56.5	271	3	US-09-077-675A-12	Sequence 12, Appl
37	35	56.5	271	3	US-09-077-674-12	Sequence 12, Appl
38	35	56.5	289	3	US-09-077-675A-10	Sequence 10, Appl
39	35	56.5	289	4	US-09-077-674-10	Sequence 10, Appl
40	35	56.5	302	3	US-09-077-675A-7	Sequence 7, Appl
41	35	56.5	302	4	US-09-077-674-7	Sequence 7, Appl
42	35	56.5	346	4	US-09-543-681A-4493	Sequence 4493, App
43	35	56.5	361	3	US-09-077-675A-8	Sequence 8, Appl
44	35	56.5	361	4	US-09-077-674-8	Sequence 13, Appl
45	35	56.5	366	3	US-09-077-675A-13	

ALIGNMENTS

RESULT 1
US-09-641-803-4
; Sequence 4, Application US/09641803
; Patent No. 6500798
; GENERAL INFORMATION:
; APPLICANT: STANTON, G. John
; APPLICANT: HUGHES, Thomas K.
; APPLICANT: BOLDOGH, Istvan
; TITLE OF INVENTION: USE OF COLOSTRIN, CONSTITUENT PEPTIDES THEREOF, AND
; TITLE OF INVENTION: ANALOGS THEREOF AS OXIDATIVE STRESS REGULATORS
; FILE REFERENCE: 265.00220101
; CURRENT APPLICATION NUMBER: US/09/641.803
; PRIOR FILING DATE: 2000-08-17
; PRIOR FILING DATE: 1999-08-17
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: peptide
US-09-641-803-4

Query Match 100.0%; Score 62; DB 4; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.00044;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qv 1 LFFFLPVNVLP 12
| | | | | | | | | | | |
Db 1 LFFFLPVNVLP 12

RESULT 2
US-09-107-532A-5890
; Sequence 5890, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCE: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA

```
;
;
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
;
; INFORMATION FOR SEQ ID NO: 5890:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 104 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEetical: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...104
; SEQUENCE DESCRIPTION: SEQ ID NO: 5890:
US-09-107-532A-5890

Query Match 66.1%; Score 41; DB 4; Length 104;
Best Local Similarity 66.7%; Pred. No. 11;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 2 FFFFLPVNV 10
Db 91 FFFFIPLNV 99

;
;
; RESULT 3
; US-09-270-767-35554
; Sequence 3554, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 35554
; LENGTH: 235
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-35554

Query Match 64.5%; Score 40; DB 4; Length 235;
Best Local Similarity 50.0%; Pred. No. 37;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LFFFLPVNV 10
Db 23 LFFFIPLNI 32

;
;
; RESULT 4
; US-09-270-767-50771
; Sequence 50771, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 50771
; LENGTH: 235
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-50771

Query Match 64.5%; Score 40; DB 4; Length 235;
Best Local Similarity 50.0%; Pred. No. 37;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LFFFLPVNV 10
Db 23 LFFFIPLNI 32

;
;
; RESULT 5
; US-09-328-352-6242
; Sequence 6242, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 6242
; LENGTH: 733
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
; OTHER INFORMATION:
US-09-328-352-6242

Query Match 64.5%; Score 40; DB 4; Length 733;
Best Local Similarity 41.7%; Pred. No. 1.2e+02;
Matches 5; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LFFFLPVNVLP 12
Db 340 ILFFVPLNMIP 351

;
;
; RESULT 6
; US-09-252-991A-23368
; Sequence 23368, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
```

```
; SEQ ID NO 23368
; LENGTH: 528
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-23368

Query Match      62.9%; Score 39; DB 4; Length 528;
Best Local Similarity 50.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Oy 2 FFFLPVNVNL 11
   |||:|:|:|
Db 395 FFFPILSIL 404

RESULT 7
US-09-489-039A-7429
; Sequence 7429, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; PRIOR FILING DATE: 2000-01-27
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 7429
; LENGTH: 483
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-7429

Query Match      61.3%; Score 38; DB 4; Length 483;
Best Local Similarity 60.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Oy 2 FFFLPVNVNL 11
   : |||:|:|
Db 130 YLFLPMLNVL 139

RESULT 8
US-09-205-258-456
; Sequence 456, Application US/09205258
; Patent No. 6525174
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 207 Human Secreted Proteins
; FILE REFERENCE: P2007E1
; CURRENT APPLICATION NUMBER: US/09/205,258
; CURRENT FILING DATE: 1998-12-04
; EARLIER APPLICATION NUMBER: PCT/US98/11422
; EARLIER FILING DATE: 1998-06-04
; EARLIER APPLICATION NUMBER: 60/048,885
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,375
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,881
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,880
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,896
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,020
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,876
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,895
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,884

; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,894
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,971
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,882
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,899
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,893
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,900
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,901
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,892
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,915
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,019
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,970
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,972
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,916
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,373
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,875
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,374
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,917
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,949
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,883
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,897
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,898
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,962
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,963
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,877
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,878
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/070,923
; EARLIER FILING DATE: 1997-12-18
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; NUMBER OF SEQ ID NOS: 1227
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 456
; LENGTH: 42
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (42)
; OTHER INFORMATION: Xaa equals stop translation
US-09-205-258-456
```

Query Match 59.7%; Score 37; DB 4; Length 42;
Best Local Similarity 75.0%; Pred. No. 19;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 LFFFLPVV 8

DB 21 LFFFLPLI 28

RESULT 9

US-09-107-532A-6843

; Sequence 6843, Application US/09107532A

; Patent No. 6583275

GENERAL INFORMATION:

APPLICANT: Lynn A Doucette-Stamm and David Bush

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS

NUMBER OF SEQUENCES: 7310

CORRESPONDENCE ADDRESS:

ADDRESSEE: GENOME THERAPEUTICS CORPORATION

STREET: 100 Beaver Street

CITY: Waltham

STATE: Massachusetts

COUNTRY: USA

ZIP: 02354

COMPUTER READABLE FORM:

MEDIUM TYPE: CD-ROM ISO9660

COMPUTER: PC

OPERATING SYSTEM: <Unknown>

SOFTWARE: ASCII

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/107,532A

FILING DATE: 30-Jun-1998

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/085,598

FILING DATE: 14 May 1998

APPLICATION NUMBER: 60/051571

FILING DATE: July 2, 1997

ATTORNEY/AGENT INFORMATION:

NAME: Ariniello, Pamela Deneke

REGISTRATION NUMBER: 40,489

REFERENCE/DOCKET NUMBER: GTC-012

TELECOMMUNICATION INFORMATION:

TELEPHONE: (781)893-5007

TELEFAX: (781)893-8277

INFORMATION FOR SEQ ID NO: 6843:

SEQUENCE CHARACTERISTICS:

LENGTH: 72 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: YES

ORIGINAL SOURCE:

ORGANISM: Enterococcus faecium

FEATURE:

NAME/KEY: misc feature

LOCATION: (3) LOCATION 1...72

SEQUENCE DESCRIPTION: SEQ ID NO: 6843:

US-09-107-532A-6843

Query Match 59.7%; Score 37; DB 4; Length 72;
Best Local Similarity 54.5%; Pred. No. 33;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 FFFFLPVNVLP 12

DB 45 FFFFLPVNVLP 55

RESULT 10

US-09-205-258-251

; Sequence 251, Application US/09205258

; Patent No. 6525174

GENERAL INFORMATION:
APPLICANT: Young et al.
TITLE OF INVENTION: 207 Human Secreted Proteins
FILE REFERENCE: PZ007P1
CURRENT APPLICATION NUMBER: US/09/205,258
CURRENT FILING DATE: 1998-12-04
EARLIER APPLICATION NUMBER: PCT/US98/11422
EARLIER FILING DATE: 1998-06-04
EARLIER APPLICATION NUMBER: 60/048,885
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,375
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,881
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,880
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,896
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,020
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,876
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,895
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,884
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,894
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,971
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,964
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,882
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,899
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,893
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,900
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,901
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,892
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,915
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,019
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,970
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,972
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,916
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,373
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,875
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,374
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,917
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,949
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,974
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,883
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,897
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,898
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,962

EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,963
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,877
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,878
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/070,923
EARLIER FILING DATE: 1997-12-18
EARLIER APPLICATION NUMBER: 60/092,921
EARLIER FILING DATE: 1998-07-15
EARLIER APPLICATION NUMBER: 60/094,657
EARLIER FILING DATE: 1998-07-30
NUMBER OF SEQ ID NOS: 1227
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 251
LENGTH: 103
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (103)
OTHER INFORMATION: Xaa equals stop translation
US-09-205-258-251

Query Match 59.7%; Score 37; DB 4; Length 103;
Best Local Similarity 75.0%; Pred. No. 48;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LFFFLPVV 8
Db 82 LFFFLPLI 89

RESULT 11

US-09-489-039A-12853
Sequence 12853, Application US/09489039A
Patent No. 6610836
GENERAL INFORMATION:
APPLICANT: Gary Breton et. al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
FILE REFERENCE: 2709.2004001
CURRENT APPLICATION NUMBER: US/09/489,039A
CURRENT FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: US 60/117,747
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 14342
SEQ ID NO 12853
LENGTH: 405
TYPE: PRT
ORGANISM: Klebsiella pneumoniae
US-09-489-039A-12853

Query Match 59.7%; Score 37; DB 4; Length 405;
Best Local Similarity 70.0%; Pred. No. 2e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 FFLPVVNVLP 12
Db 303 FFLPVVNVLP 312

RESULT 12

US-09-071-035-196
Sequence 196, Application US/09071035
Patent No. 6448043
GENERAL INFORMATION:
APPLICANT: Gil H. Choi
TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
NUMBER OF SEQUENCES: 496
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/071,035
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: A. Anders Brookes
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB369P2
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 196:
SEQUENCE CHARACTERISTICS:
LENGTH: 48 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-071-035-196

Query Match 58.1%; Score 36; DB 4; Length 48;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 PVNVLP 12
Db 24 PVNVLP 30

RESULT 13

US-09-071-035-194
Sequence 194, Application US/09071035
Patent No. 6448043
GENERAL INFORMATION:
APPLICANT: Gil H. Choi
TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
NUMBER OF SEQUENCES: 496
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/071,035
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: A. Anders Brookes
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB369P2

TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 194:
SEQUENCE CHARACTERISTICS:
LENGTH: 109 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-071-035-194

Query Match 58.1%; Score 36; DB 4; Length 109;
Best Local Similarity 100.0%; Pred. No. 74;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 PVNVLP 12
Db 48 PVNVLP 54

RESULT 14

US-09-134-000C-4445
Sequence 4445, Application US/09134000C

Patent No. 6617156

GENERAL INFORMATION:

APPLICANT: Lynn Doucette-Stamm et al

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO

FILE REFERENCE: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: 032796-032

CURRENT APPLICATION NUMBER: US/09/134,000C

CURRENT FILING DATE: 1998-08-13

PRIOR APPLICATION NUMBER: US 60/055,778

PRIOR FILING DATE: 1997-08-15

NUMBER OF SEQ ID NOS: 6812

SOFTWARE: PatentIn version 3.1

SEQ ID NO 4445

LENGTH: 123

TYPE: PRT

ORGANISM: Enterococcus faecalis

US-09-134-000C-4445

Query Match 58.1%; Score 36; DB 4; Length 123;

Best Local Similarity 100.0%; Pred. No. 84;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 PVNVLP 12
Db 62 PVNVLP 68

RESULT 15

US-09-489-039A-11180

Sequence 11180, Application US/09489039A

Patent No. 6610836

GENERAL INFORMATION:

APPLICANT: Gary Breton et. al

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA

FILE REFERENCE: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: 2709.2004001

CURRENT APPLICATION NUMBER: US/09/489,039A

CURRENT FILING DATE: 2000-01-27

PRIOR APPLICATION NUMBER: US 60/117,747

PRIOR FILING DATE: 1999-01-29

NUMBER OF SEQ ID NOS: 14342

SEQ ID NO 11180

LENGTH: 137

TYPE: PRT

ORGANISM: Klebsiella pneumoniae

US-09-489-039A-11180

Query Match 58.1%; Score 36; DB 4; Length 137;

Best Local Similarity 45.5%; Pred. No. 94;

Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LFEFLPVNVLP 11

Db 123 LFFPIPVLSII 133

Search completed: April 12, 2005, 08:07:51
Job time : 15.0727 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 12, 2005, 07:49:30 ; Search time 37.2 Seconds
(without alignments)
107.096 Million cell updates/sec

Title: US-10-691-330-4
Perfect score: 62
Sequence: 1 LFFFLPVVNVLP 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1418010 seqs, 33197259 residues

Total number of hits satisfying chosen parameters: 1418010

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	62	100.0	12	14 US-10-281-652-4	Sequence 4, Appli
2	62	100.0	12	17 US-10-691-330-4	Sequence 4, Appli
3	45	72.6	96	15 US-10-424-599-261912	Sequence 211912,
4	44	71.0	47	15 US-10-424-599-211893	Sequence 211893,
5	43	69.4	54	15 US-10-424-599-143141	Sequence 143141,
6	43	69.4	180	15 US-10-424-599-227817	Sequence 227817,
7	43	69.4	197	15 US-10-424-599-218474	Sequence 218474,
8	43	69.4	678	15 US-10-424-599-260138	Sequence 260138,
9	42	67.7	72	15 US-10-424-599-197974	Sequence 197974,
10	42	67.7	101	15 US-10-424-599-245079	Sequence 245079,
11	41	66.1	54	15 US-10-424-599-234012	Sequence 234012,
12	41	66.1	56	15 US-10-424-599-154599	Sequence 154599,
13	41	66.1	81	15 US-10-424-599-228846	Sequence 228846,

14	41	66.1	82	15	US-10-424-599-252118	Sequence 252118,
15	41	66.1	126	15	US-10-094-749-1945	Sequence 1945, Ap
16	41	66.1	141	15	US-10-424-599-272793	Sequence 272793,
17	41	66.1	153	15	US-10-424-599-171577	Sequence 171577,
18	41	66.1	666	16	US-10-437-963-138531	Sequence 138531,
19	41	66.1	738	16	US-10-437-963-138533	Sequence 138533,
20	40	64.5	98	15	US-10-425-114-42601	Sequence 42601, A
21	40	64.5	132	16	US-10-437-963-187827	Sequence 187827,
22	40	64.5	363	16	US-10-437-963-167680	Sequence 167680,
23	39	62.9	46	15	US-10-424-599-149052	Sequence 149052,
24	39	62.9	53	15	US-10-424-599-208834	Sequence 208834,
25	39	62.9	58	15	US-10-424-599-181046	Sequence 181046,
26	39	62.9	95	15	US-10-424-599-237727	Sequence 237727,
27	39	62.9	118	15	US-10-424-599-164024	Sequence 164024,
28	39	62.9	356	16	US-10-437-963-167720	Sequence 167720,
29	39	62.9	422	15	US-10-282-122A-49461	Sequence 49461, A
30	38	61.3	68	15	US-10-424-599-204168	Sequence 204168,
31	38	61.3	68	16	US-10-437-963-140548	Sequence 140548,
32	38	61.3	85	15	US-10-424-599-244849	Sequence 244849,
33	38	61.3	108	16	US-10-437-963-192194	Sequence 192194,
34	38	61.3	563	16	US-10-437-963-107198	Sequence 107198,
35	37	59.7	42	10	US-09-933-767-456	Sequence 456, App
36	37	59.7	42	14	US-10-004-860-456	Sequence 456, App
37	37	59.7	42	14	US-10-023-282-456	Sequence 199617,
38	37	59.7	52	16	US-10-437-963-199617	Sequence 235499,
39	37	59.7	53	15	US-10-424-599-235499	Sequence 932, App
40	37	59.7	63	14	US-10-083-357-932	Sequence 256914,
41	37	59.7	72	15	US-10-424-599-255914	Sequence 44750, A
42	37	59.7	81	16	US-10-767-701-44750	Sequence 1107, Ap
43	37	59.7	84	9	US-09-764-877-1107	Sequence 1107, Ap
44	37	59.7	84	15	US-10-242-515-1107	Sequence 109671,
45	37	59.7	89	16	US-10-437-963-109671	

ALIGNMENTS

RESULT 1

US-10-281-652-4
; Sequence 4, Application US/10281652
; Publication No. US20030091606A1
; GENERAL INFORMATION:
; APPLICANT: STANTON, G. John
; APPLICANT: HUGHES, Thomas K.
; APPLICANT: BOLDOGH, Istvan
; TITLE OF INVENTION: USE OF COLOSTRININ, CONSTITUENT PEPTIDES THEREOF, AND
; FILE REFERENCE: 265.00220101
; CURRENT APPLICATION NUMBER: US/10/281.652
; CURRENT FILING DATE: 2002-10-28
; PRIOR FILING DATE: 2000-08-17
; PRIOR APPLICATION NUMBER: 60/149,310
; PRIOR FILING DATE: 1999-08-17
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: peptide
US-10-281-652-4

Query Match 100.0%; Score 62; DB 14; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.0008;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 LFFFLPVVNVLP 12
Db 1 LFFFLPVVNVLP 12

```
RESULT 2
US-10-691-330-4
; Sequence 4, Application US/10691330
; Publication No. US20050042300A1
; GENERAL INFORMATION:
; APPLICANT: Boldogh, Istvan
; APPLICANT: Stanton, G. John
; APPLICANT: Georgiades, Jerzy A.
; APPLICANT: Hughes, Thomas K., Jr.
; APPLICANT: Kruzel, Marian
; TITLE OF INVENTION: USE OF COLOSTRININ, CONSTITUENT PEPTIDES THEREOF, AND ANALOGS
; TITLE OF INVENTION: THEREOF AS INHIBITORS OF APOPTOSIS AND OTHER CELLULAR DAMAGE
; FILE REFERENCE: 265.00390101
; CURRENT APPLICATION NUMBER: US/10/691.330
; CURRENT FILING DATE: 2003-10-22
; PRIOR FILING DATE: 2002-10-22
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4
; LENGTH: 12
; TYPE: PRT
; ORGANISM: artificial
; FEATURE:
; OTHER INFORMATION: Synthetic Peptides
US-10-691-330-4

Query Match      100.0%; Score 62; DB 17; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.0008;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LFFFLPVNVLP 12
Db 1 LFFFLPVNVLP 12

RESULT 3
US-10-424-599-261912
; Sequence 261912, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424.599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 261912
; LENGTH: 96
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_78529C.1.pep
US-10-424-599-261912

Query Match      72.6%; Score 45; DB 15; Length 96;
Best Local Similarity 72.7%; Pred. No. 4.9;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 FFFFLPVNVLP 12
Db 12 FFFFLPVNVLP 22

RESULT 4
US-10-424-599-211893
; Sequence 211893, Application US/10424599
; Publication No. US20040031072A1
```

```
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424.599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 211893
; LENGTH: 47
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_33366C.1.pep
US-10-424-599-211893

Query Match      71.0%; Score 44; DB 15; Length 47;
Best Local Similarity 66.7%; Pred. No. 3.4;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 LFFFLPVNVLP 12
Db 34 LFFFLDKINLLP 45

RESULT 5
US-10-424-599-143141
; Sequence 143141, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424.599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 143141
; LENGTH: 54
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_10026C.1.pep
US-10-424-599-143141

Query Match      69.4%; Score 43; DB 15; Length 54;
Best Local Similarity 63.6%; Pred. No. 5.8;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 2 FFFFLPVNVLP 12
Db 39 FFFLPITNALP 49

RESULT 6
US-10-424-599-227817
; Sequence 227817, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424.599
```

; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 227817
; LENGTH: 180
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_47747C.1.pep
US-10-424-599-227817

Query Match 69.4%; Score 43; DB 15; Length 180;
Best Local Similarity 77.8%; Pred. No. 20;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LFFFLPVVN 9
|||:||||:
Db 125 LFFFLPVVN 133

RESULT 7

US-10-424-599-218474
; Sequence 218474, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei

; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 218474
; LENGTH: 197
; TYPE: PRT

; ORGANISM: Glycine max

; FEATURE:

; NAME/KEY: unsure

; LOCATION: (1)..(197)

; OTHER INFORMATION: unsure at all Xaa locations

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT_MRT3847_3930C.1.pep

US-10-424-599-218474

Query Match 69.4%; Score 43; DB 15; Length 197;
Best Local Similarity 77.8%; Pred. No. 22;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LFFFLPVVN 9
|||:||||:
Db 126 LFFFLPVVN 134

RESULT 8

US-10-424-599-260138
; Sequence 260138, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei

; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 260138
; LENGTH: 678
; TYPE: PRT

; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(678)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_76929C.1.pep
US-10-424-599-260138

Query Match 69.4%; Score 43; DB 15; Length 678;
Best Local Similarity 77.8%; Pred. No. 80;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LFFFLPVVN 9
|||:||||:
Db 74 LFFFLPVVN 82

RESULT 9

US-10-424-599-197974
; Sequence 197974, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei

; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 197974
; LENGTH: 72
; TYPE: PRT

; ORGANISM: Glycine max

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT_MRT3847_20796C.1.pep

US-10-424-599-197974

Query Match 67.7%; Score 42; DB 15; Length 72;
Best Local Similarity 63.6%; Pred. No. 11;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 FFFLPVVNVLP 12
|||:|:|:
Db 15 FFFLEINSLP 25

RESULT 10

US-10-424-599-245079
; Sequence 245079, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei

; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 245079
; LENGTH: 101
; TYPE: PRT

; ORGANISM: Glycine max

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT_MRT3847_63337C.1.pep

US-10-424-599-245079

Query Match 67.7%; Score 42; DB 15; Length 101;
Best Local Similarity 87.5%; Pred. No. 16;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 FFFLPVNV 9
 |||||:
DB 76 FFFLPVNV 83

RESULT 11
US-10-424-599-234012
; Sequence 234012, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 234012
; LENGTH: 54
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_53337C.1.pep
US-10-424-599-234012

Query Match 66.1%; Score 41; DB 15; Length 54;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LFFFLPVV 8
 |||||:
DB 8 LFFFLPVV 15

RESULT 12
US-10-424-599-154599
; Sequence 154599, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 154599
; LENGTH: 56
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_110624C.1.pep
US-10-424-599-154599

Query Match 66.1%; Score 41; DB 15; Length 56;
Best Local Similarity 63.6%; Pred. No. 13;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 LFFFLPVNV 11
 |||||:
DB 18 LFFFLPIS 28

RESULT 13

US-10-424-599-228846
; Sequence 228846, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 228846
; LENGTH: 81
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_48675C.1.pep
US-10-424-599-228846

Query Match 66.1%; Score 41; DB 15; Length 81;
Best Local Similarity 71.4%; Pred. No. 19;
Matches 10; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

QY 1 LFFFLPV--VNVLP 12
 |||||:
DB 54 LFFFLPVHVNVLP 67

RESULT 14

US-10-424-599-252118
; Sequence 252118, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 252118
; LENGTH: 82
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_69690C.1.pep
US-10-424-599-252118

Query Match 66.1%; Score 41; DB 15; Length 82;
Best Local Similarity 60.0%; Pred. No. 19;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 FFLPVNVLP 12
 |:|:|:
DB 11 FLVPIVNVLP 20

RESULT 15

US-10-094-749-1945
; Sequence 1945, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI

; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEXI, NAOHICO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTUYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
; FILE REFERENCE: 084335/0160
; CURRENT APPLICATION NUMBER: US/10/094,749
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 1945
; LENGTH: 126
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-094-749-1945

Query Match 66.1%; Score 41; DB 15; Length 126;
Best Local Similarity 72.7%; Pred. No. 30;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 FFFLPVNVLP 12
||| |:
Db 35 FFFLPPVSSLP 45

Search completed: April 12, 2005, 07:55:21
Job time : 38.2 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 12, 2005, 07:49:31 ; Search time 8.61818 Seconds
(without alignments)
133.973 Million cell updates/sec

Title: US-10-691-330-4

Perfect score: 62

Sequence: 1 LFFFLPVNVLP 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_79:*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	48	77.4	469	2 AB3519	enterobactin synth
2	45	72.6	144	2 D90125	hypothetical prote
3	44	71.0	62	2 A12652	hypothetical prote
4	41	66.1	66	2 S49114	hypothetical prote
5	41	66.1	488	2 D70420	NADH2 dehydrogenas
6	41	66.1	692	2 S61200	probable membrane
7	39	62.9	256	2 I49846	spa29 protein - Sh
8	39	62.9	359	2 T1247	hypothetical prote
9	39	62.9	391	2 S74688	hypothetical prote
10	39	62.9	471	2 F72257	lipopolysaccharide
11	39	62.9	509	2 E93002	drug efflux transp
12	39	62.9	537	2 G31277	quinate transport
13	39	62.9	710	2 T52573	cyclic nucleotide
14	39.5	62.1	364	2 AD2302	hypothetical prote
15	38	61.3	221	2 AF3483	heme exporter prot
16	38	61.3	296	2 D98345	SN-glycerol-3-phos
17	38	61.3	296	2 AC2937	hypothetical 3-phos
18	38	61.3	461	2 H85755	probable amino aci
19	38	61.3	479	2 C84878	probable amino aci
20	38	61.3	479	2 A90863	probable amino aci
21	38	61.3	532	2 G82872	cardiolipin syntha
22	38	61.3	678	2 G84682	hypothetical prote
23	38	61.3	738	2 E86294	hypothetical prote
24	38	61.3	746	2 H86330	probable cyclic nu
25	37	59.7	218	2 T15420	hypothetical prote
26	37	59.7	263	2 AD0851	secretory protein
27	37	59.7	263	2 S37309	spa9 protein - Sal
28	37	59.7	546	2 G71348	probable apolipop
29	37	59.7	851	2 T31520	hypothetical prote

30 36 58.1 118 2 S76608
31 36 58.1 201 2 T33593
32 36 58.1 249 2 G64470
33 36 58.1 252 2 A69286
34 36 58.1 302 2 E72402
35 36 58.1 306 1 CPBOB
36 36 58.1 310 2 AB0979
37 36 58.1 310 2 F90011
38 36 58.1 327 1 HLHUCD
39 36 58.1 415 2 A32129
40 36 58.1 438 2 H82280
41 36 58.1 488 2 T30602
42 36 58.1 501 2 T52170
43 36 58.1 505 2 T02898
44 36 58.1 555 2 B72341
45 36 58.1 689 2 D85013

ALIGNMENTS

RESULT 1

AB3519

enterobactin synthetase component F [imported] - Brucella melitensis (strain 16M)
C:Species: Brucella melitensis

C>Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004

C:Accession: AB3519

R:DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,

; Mazur, M.; Goitsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess

Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002

A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis

A:Reference number: AD3252; PMID:11756688

A:Accession: AB3519

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-469 <KUR>

A:Cross-references: UNIPROT:Q8YDU8; GB:AE0808918; PIDN:AA153317.1; PID:gl7984203; GSPDB:G

A:Experimental source: strain 16M

C:Genetics:

A:Gene: BME110076

A:Map position: 11

Query Match 77.4%; Score 48; DB 2; Length 469;
Best Local Similarity 72.7%; Pred. No. 1;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 FFFLPVNVLP 12

Db 374 FFFSPLINVLP 384

RESULT 2

D90125

hypothetical protein orf144 [imported] - Guillardia theta nucleomorph

C:Species: nucleomorph Guillardia theta

A:Note: a nucleomorph is the vestigial nucleus of a eukaryotic endosymbiont

C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004

C:Accession: D90125

R:Douglas, S.; Zauner, S.; Fraunholz, M.; Beaton, M.; Penny, S.; Deng, L.T.; Wu, X.; Rei

Nature 410, 1091-1096, 2001

A:Title: The highly reduced genome of an enslaved algal nucleus.

A:Reference number: A99082; MUID:11323671; PMID:11323671

A:Accession: D90125

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-144 <DOU>

A:Cross-references: UNIPROT:Q98S86; GB:AF083031; NID:gl13794319; PIDN:AAK39696.1; GSPDB:G

C:Genetics:

A:Gene: orf144

A:Map position: 3

A:Genome: nucleomorph

C:Keywords: nucleomorph

Query Match 72.6%; Score 45; DB 2; Length 144;
 Best Local Similarity 66.7%; Pred. No. 1.1; Mismatches 2; Indels 0; Gaps 0;
 Matches 8; Conservative 24

QY 1 LFFFLPVNVLP 12
 :|||:|:|
 DB 24 IFFFLKVNILP 35
 :|||:|:|

RESULT 3
 A12652
 hypothetical protein Atu0623 [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
 C:Species: Agrobacterium tumefaciens
 C>Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
 C:Accession: A12652
 R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.B.; Chen, Y.; Woo, I.;
 erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell
 ; Karp, P.; Romero, P.; Zhang, S.
 Science 294, 2317-2323, 2001
 A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
 ster, E.W.
 A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
 A:Reference number: AB2577; MUID:21608550; PMID:11743193
 A:Accession: A12652
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-62 <KUR>
 A:Cross-references: UNIPROT:Q8UHQ7; GB:AE008688; PIDN:AAL41639.1; PID:g17738979; GSPDB:C
 A:Experimental source: strain C58 (Dupont)
 C:Genetics:
 A:Gene: Atu0623
 A:Map position: circular chromosome

Query Match 71.0%; Score 44; DB 2; Length 62;
 Best Local Similarity 72.7%; Pred. No. 0.69; Mismatches 2; Indels 0; Gaps 0;
 Matches 8; Conservative 5

QY 2 FFFFLPVNVLP 12
 :|||:|:|
 DB 5 FVSLPVNVLP 15
 :|||:|:|

RESULT 4
 S49114
 hypothetical protein - yeast (Williopsis suaveolens) mitochondrion (fragment)
 C:Species: mitochondrion Williopsis suaveolens
 C>Date: 16-Feb-1995 #sequence_revision 26-May-1995 #text_change 09-Jul-2004
 C:Accession: S49114
 R:Nosek, J.
 submitted to the EMBL Data Library, January 1994
 A:Reference number: S49114
 A:Accession: S49114
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-66 <NOS>
 A:Cross-references: UNIPROT:Q36231; EMBL:X77238; NID:G509359; PID:G509360
 C:Genetics:
 A:Genome: mitochondrion
 A:Genetic code: SGC2
 C:Keywords: mitochondrion

Query Match 66.1%; Score 41; DB 2; Length 66;
 Best Local Similarity 50.0%; Pred. No. 2.4; Mismatches 4; Indels 0; Gaps 0;
 Matches 6; Conservative 37

QY 1 LFFFLPVNVLP 12
 :|||:|:|
 DB 37 LFFFLIIGVMP 48
 :|||:|:|

RESULT 5
 D70420
 NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) I chain nuoN2 - Aquifex aeolicus

C:Species: Aquifex aeolicus
 C>Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 09-Jul-2004
 C:Accession: D70420
 R:Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Ove
 V.
 Nature 392, 353-358, 1998
 A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
 A:Reference number: A70300; MUID:98196666; PMID:9537320
 A:Accession: D70420
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-488 <AQF>
 A:Cross-references: UNIPROT:O67391; GB:AE000737; NID:G2983782; PIDN:AAC07354.1; PID:G298
 A:Experimental source: strain VFS
 C:Genetics:
 A:Gene: nuoN2
 C:Superfamily: NADH dehydrogenase (ubiquinone) chain 2
 C:Keywords: membrane-associated complex; NAD; oxidoreductase

Query Match 66.1%; Score 41; DB 2; Length 488;
 Best Local Similarity 58.3%; Pred. No. 18; Mismatches 7; Conservative 3; Indels 0; Gaps 0;
 Matches 255

QY 1 LFFFLPVNVLP 12
 :|||:|:|
 DB 255 LFFFLPVNVLP 266
 :|||:|:|

RESULT 6
 S61200
 probable membrane protein YDR314c - yeast (Saccharomyces cerevisiae)
 N:Alternate names: hypothetical protein D9740.21
 C:Species: Saccharomyces cerevisiae
 C>Date: 23-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 09-Jul-2004
 C:Accession: S61200
 R:Ding, H.
 submitted to the EMBL Data Library, June 1995
 A:Description: The sequence of S. cerevisiae cosmid 9740.
 A:Reference number: S61160
 A:Accession: S61200
 A:Molecule type: DNA
 A:Residues: 1-692 <DIN>
 A:Cross-references: UNIPROT:Q06665; EMBL:U28374; NID:G849207; PIDN:AAB64750.1; PID:G84920

C:Genetics:
 A:Gene: MIP5.YDR314C
 A:Cross-references: SGD:S0002722
 A:Map position: 4R
 C:Superfamily: yeast probable membrane protein YDR314c
 C:Keywords: transmembrane protein
 F;94-110/Domain: transmembrane #status predicted <TM1>
 F;239-255/Domain: transmembrane #status predicted <TM2>

Query Match 66.1%; Score 41; DB 2; Length 692;
 Best Local Similarity 66.7%; Pred. No. 25; Mismatches 8; Conservative 2; Indels 0; Gaps 0;
 Matches 242

QY 1 LFFFLPVNVLP 12
 :|||:|:|
 DB 242 LFFFLIENVLP 253
 :|||:|:|

RESULT 7
 I49846
 spa29 protein - Shigella flexneri plasmid pMYSH6000
 C:Species: Shigella flexneri
 C>Date: 12-May-1994 #sequence_revision 12-May-1994 #text_change 09-Jul-2004
 C:Accession: I49846
 R:Sasakawa, C.; Komatsu, K.; Tobe, T.; Suzuki, T.; Yoshikawa, M.
 J. Bacteriol. 175, 2334-2346, 1993
 A:Title: Eight genes in region 5 that form an operon are essential for invasion of epithe
 A:Reference number: A49846; MUID:93224456; PMID:8385666
 A:Accession: I49846
 A:Status: preliminary

A:Molecule type: DNA
A:Residues: 1-256 <SAS>
A:Cross-references: UNIPROT:P40706; GB:D13663; NID:g287439; PIDN:BAA02831.1; PID:g303895
C:Genetics:
A:Genome: plasmid
C:Superfamily: Shigella flexneri spa29 protein
C:Keywords: transmembrane protein

Query Match 62.9%; Score 39; DB 2; Length 256;
Best Local Similarity 77.8%; Pred. No. 21;
Matches 7; Conservative 1; Mismatches 0; Gaps 0;

Qy 1 LFFFLPVVN 9
| | | | | : |
Db 27 LFFFLPFLN 35

RESULT 8
T21247
hypothetical protein F22B8.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T21247
R:McMurray, A.
A:Reference number: 219396
A:Accession: T21247
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-359 <WIL>
A:Cross-references: UNIPROT:O17830; EMBL:Z83106; PIDN:CAB05493.1; GSPDB:GN00023; CESP:F22B8
A:Experimental source: clone F22B8
C:Genetics:
A:Gene: CESP:F22B8.1
A:Map position: 5
A:Introns: 96/3; 132/2; 161/3; 200/3

Query Match 62.9%; Score 39; DB 2; Length 359;
Best Local Similarity 70.0%; Pred. No. 29;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 LFFFLPVVN 10
| | | | | : |
Db 312 LFFFLPIGV 321

RESULT 9
S74688
hypothetical protein sll1200 - Synecocystis sp. (strain PCC 6803)
C:Species: Synecocystis sp.
A:Variety: PCC 6803
C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C:Accession: S74688
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.; O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda, N. K.; 109-136, 1996
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synecocystis sp.

A:Reference number: S74322; MUID:97061201; PMID:8905231
A:Accession: S74688
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-391 <KAN>
A:Cross-references: UNIPROT:P72824; EMBL:D90801; GB:AB001339; NID:g1651897; PIDN:BAAL683
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 62.9%; Score 39; DB 2; Length 391;
Best Local Similarity 70.0%; Pred. No. 31;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 LFFFLPVVN 10
| | | | | : |
Db 99 LVFFFLPIANV 108

RESULT 10

F72257
lipopolysaccharide biosynthesis protein-related protein - Thermotoga maritima (strain MS)
C:Species: Thermotoga maritima
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
C:Accession: F72257
R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey, Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.; C.M.
Nature 399, 323-329, 1999
A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome sequencing
A:Reference number: A72200; MUID:99287316; PMID:10360571
A:Accession: F72257
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-471 <ARN>
A:Cross-references: UNIPROT:Q9X1C5; GB:AE001793; GB:AE000512; NID:g4981963; PIDN:AAD3647
A:Experimental source: strain MSB8
C:Genetics:
A:Gene: TM1405

Query Match 62.9%; Score 39; DB 2; Length 471;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 FLPVVNVL 11
| | | | |
Db 153 FLPVVNVL 160

RESULT 11

E83002
drug efflux transporter PAS160 [imported] - Pseudomonas aeruginosa (strain PAO1)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Br. adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, L.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic pathogen
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: E83002
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-509 <STO>
A:Cross-references: UNIPROT:O85163; GB:AE004928; GB:AE0004091; NID:g9951450; PIDN:AAG0854
A:Experimental source: strain PAO1
C:Genetics:
A:Gene: PAS160
C:Superfamily: lincomycin-resistance protein lmrB

Query Match 62.9%; Score 39; DB 2; Length 509;
Best Local Similarity 50.0%; Pred. No. 41;
Matches 5; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy 2 FFLPVNVN 11
| | | | : |
Db 376 FFFPILSIL 385

RESULT 12

G31277
guinate transport protein - Neurospora crassa (tentative sequence)
N:Alternate names: guinate transporter
C:Species: Neurospora crassa
C:Date: 26-Apr-1989 #sequence_revision 26-Apr-1989 #text_change 09-Jul-2004
C:Accession: S04254; G31277
R:Geever, R.F.; Huet, L.; Baum, J.A.; Tyler, B.M.; Patel, V.B.; Rutledge, B.J.; Case, M. J. Mol. Biol. 207, 15-34, 1989
A:Title: DNA sequence, organization and regulation of the qa gene cluster of Neurospora

A;Reference number: S04250; MUID:89293848; PMID:2525625
A;Accession: S04254
A;Molecule type: DNA
A;Residues: 1-537 <GE2>
A;Cross-references: UNIPROT:P11636; EMBL:X14603; NID:g3060; PIDN:CAA32752.1; PID:g3065
C;Genetics:
A;Gene: qa-Y

C;Superfamily: maltose transport protein MAL61

C;Keywords: transmembrane protein
F:22-42/Domain: transmembrane #status predicted <TM01>
F:67-87/Domain: transmembrane #status predicted <TM02>
F:99-119/Domain: transmembrane #status predicted <TM03>
F:132-152/Domain: transmembrane #status predicted <TM04>
F:161-181/Domain: transmembrane #status predicted <TM05>
F:195-215/Domain: transmembrane #status predicted <TM06>
F:286-306/Domain: transmembrane #status predicted <TM07>
F:324-344/Domain: transmembrane #status predicted <TM08>
F:356-376/Domain: transmembrane #status predicted <TM09>
F:390-410/Domain: transmembrane #status predicted <TM10>
F:427-447/Domain: transmembrane #status predicted <TM11>
F:459-479/Domain: transmembrane #status predicted <TM12>

Query Match 62.9%; Score 39; DB 2; Length 537;
Best Local Similarity 50.0%; Pred. No. 43;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 LFFFLPVVNVLP 12
Db 475 IYFFLPVTKSIP 486
:::|||||

RESULT 13

T52573
cyclic nucleotide and calmodulin-regulated ion channel [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 24-Oct-2000 #sequence_revision 24-Oct-2000 #text_change 09-Jul-2004
C;Accession: T52573
R;Kohler, C.; Merkle, T.; Neuhaus, G.

A;Title: Characterisation of a novel gene family of putative cyclic nucleotide- and calmodulin-regulated ion channels
A;Reference number: 226120
A;Accession: T52573
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-710 <KOH>
A;Cross-references: UNIPROT:Q8RWS9; EMBL:Y17913; PIDN:CAB40130.1
A;Experimental source: cultivar Columbia
C;Genetics:
A;Gene: cngc5

Query Match 62.9%; Score 39; DB 2; Length 710;
Best Local Similarity 75.0%; Pred. No. 57;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 FFFLPVVN 9
Db 110 FFYLPVIN 117
||:|||||

RESULT 14

AD2302
hypothetical protein all3971 [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C;Accession: AD2302
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yasuda, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena PCC 7120
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AD2302
A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-364 <KUR>

A;Cross-references: UNIPROT:Q8YQ64; GB:BA000019; PIDN:BAB75670.1; PID:g17133105; GSPDB:G

A;Experimental source: strain PCC 7120

C;Genetics:

A;Gene: all3971

Query Match 62.1%; Score 38.5; DB 2; Length 364;
Best Local Similarity 47.4%; Pred. No. 36;
Matches 9; Conservative 2; Mismatches 1; Indels 7; Gaps 1;

QY 1 LFFFLPVVNVLP 12
Db 281 LFFFLPVVNVLP 299
|||||

RESULT 15

AF3483

heme exporter protein B [imported] - Brucella melitensis (strain 16M)

C;Species: Brucella melitensis

C;Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004

C;Accession: AF3483

R;DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova, N.

Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002

A;Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis

A;Reference number: AD3252; PMID:11756688

A;Accession: AF3483

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-221 <KUR>

A;Cross-references: UNIPROT:Q8YEM6; GB:AE008917; PIDN:AAL53033.1; PID:g17983891; GSPDB:G

A;Experimental source: strain 16M

C;Genetics:

A;Gene: BME11852

A;Map position: I

C;Superfamily: cytochrome c biogenesis protein CycW

Query Match 61.3%; Score 38; DB 2; Length 221;
Best Local Similarity 50.0%; Pred. No. 27;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 LFFFLPVVNVLP 12
Db 23 ILFFFLAVISMP 34
||:|||||

Search completed: April 12, 2005, 07:56:51
Job time : 10.6182 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 12, 2005, 07:49:31 ; Search time 56.9454 Seconds
(without alignments)
107.909 Million cell updates/sec

Title: US-10-691-330-4
Perfect score: 62
Sequence: 1 LFFFLPVVNVLP 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48	77.4	469	2	Q8YD8
2	48	77.4	469	2	Q8FXP8
3	45	72.6	144	2	Q98S86
4	44	71.0	62	2	Q8UHQ7
5	44	71.0	218	2	Q9G873
6	44	71.0	249	2	O24102
7	44	71.0	271	2	Q8FUV9
8	44	71.0	288	2	Q8VOL7
9	44	71.0	525	2	Q8HIU9
10	42	67.7	545	2	Q8A411
11	42	67.7	585	2	Q7M9X9
12	41	66.1	66	2	Q36231
13	41	66.1	126	2	Q96NK3
14	41	66.1	242	2	Q65223
15	41	66.1	376	2	Q8IM69
16	41	66.1	488	2	Q67391
17	41	66.1	692	2	Q06665
18	41	66.1	694	2	Q65225
19	41	66.1	1085	2	Q7R3H1
20	40	64.5	85	2	Q6K9X2
21	40	64.5	352	2	Q8PUB3
22	40	64.5	431	2	Q998L1
23	40	64.5	455	2	Q73QU3
24	40	64.5	458	2	Q9XWU8
25	40	64.5	471	2	Q8EYGG
26	40	64.5	675	2	Q6ZG24
27	40	64.5	724	2	Q7X641
28	40	64.5	1441	2	Q89TS8
29	39	62.9	62	2	Q819H8
30	39	62.9	104	2	Q8PVK3
31	39	62.9	222	2	Q6Jt28

32 39 62.9 256 1 SPAR SHIFL P40706 shigella fl
33 39 62.9 305 2 Q88GD7 Q88gd7 pseudomonas
34 39 62.9 308 2 Q697F7 Q697f7 neomaskelli
35 39 62.9 359 2 O17830 O17830 caenorhabdi
36 39 62.9 331 2 P72824 P72824 synchocyst
37 39 62.9 430 2 Q8VNV9 Q8vrv9 clostridium
38 39 62.9 430 2 Q8XHA3 Q8xha3 clostridium
39 39 62.9 431 2 Q9BMP8 Q9bmp8 plasmodium
40 39 62.9 436 2 Q8KR72 Q8kr72 photorhabdu
41 39 62.9 471 2 Q9X1C5 Q9x1c5 thymotoga
42 39 62.9 499 2 Q74FI6 Q74fi6 geobacter s
43 39 62.9 509 2 Q8S163 Q8s163 pseudomonas
44 39 62.9 552 2 Q9CFI8 Q9cpi8 pasteurella
45 39 62.9 585 2 Q6FV80 Q6fv80 candida gla

ALIGNMENTS

RESULT 1

Q8YD8 PRELIMINARY; PRT; 469 AA.
AC Q8YD8; 2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE ENTEROACTIN SYNTHETASE COMPONENT F.
GN OrderedLocusNames=BMEI10076;
OS Brucella melitensis.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Brucellaceae; Brucella.
OX NCBI_TaxID=29459;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=16M / ATCC 23456 / Biotype 1;
RX MEDLINE=20020109; PubMed=11756688; DOI=10.1073/pnas.221575398;
RA DelVecchio V.G., Kapatral V., Redkar R.J., Patra G., Mufier C., Los T.,
RA Ivanova N., Anderson I., Bhattacharyya A., Lykidis A., Reznik G.,
RA Jablonski L., Larsen N., D'Souza M., Bernal A., Mazur M., Goltzman E.,
RA Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Letesson J.-J.,
RA Haeelkorn R., Kyrpides N.C., Overbeek R.;
RT "The genome sequence of the facultative intracellular pathogen
RT Brucella melitensis.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).
DR EMBL; AE009646; AAL53317.1; -.
DR PIR; AB3519; AB3519.
DR HSP; Q9KTV9; IL5A.
DR Pfam; PF00668; Condensation; 1.
KW Complete proteome.
SQ SEQUENCE 469 AA; 53074 MW; 17A7B73A02428D46 CRC64;

Query Match 77.4%; Score 48; DB 2; Length 469;
Best Local Similarity 72.7%; Pred. No. 6.8;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 FFFLPVVNVLP 12
||| :|||
Db 374 FFFSPLINVLP 384

RESULT 2

Q8FXP8 PRELIMINARY; PRT; 469 AA.
AC Q8FXP8; 2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Enterobactin synthetase, component F, putative.
GN OrderedLocusNames=BR0017;
OS Brucella suis.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Brucellaceae; Brucella.
OX NCBI_TaxID=29461;

```

RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=1330 / Biovar 1;
RX  MEDLINE=22247741; PubMed=12271122; DOI=10.1073/pnas.192319099;
RA  Paulsen I.T., Seshadri R., Nelson K.E., Eisen J.A., Heidelberg J.F.,
RA  Read T.D., Dodson R.J., Umayam L.A., Brinkac L.M., Beanan M.J.
RA  Daugherty S.C., DeBoy R.T., Durkin A.S., Kolonay J.F., Madupu R.,
RA  Nelson W.C., Ayodeji B., Kraul M., Shetty J., Malek J.A.,
RA  Van Aken S.E., Riedmuller S., Tettelin H., Gill S.R., White O.,
RA  Salzberg S.L., Hoover D.L., Lindler L.E., Halling S.M., Boyle S.M.,
RA  Fraser C.M.;
RT  "The Brucella suis genome reveals fundamental similarities between
RT  animal and plant pathogens and symbionts";
RL  Proc. Natl. Acad. Sci. U.S.A. 99:13148-13153 (2002).
DR  EMBL; AE014506; AAN33229.1; -.
DR  HSSP; Q9KTV9; 1L5A.
DR  TIGR; BRA0017; -.
DR  InterPro; IPR001242; Condensatn.
DR  Pfam; PF00668; Condensation; 1.
KW  Complete proteome.
SQ  SEQUENCE 469 AA; 52986 MW; AD46038DC9F854A31 CRC64;

Query Match 77.4%; Score 48; DB 2; Length 469;
Best Local Similarity 72.7%; Pred. No. 6.8;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY  2 PFPLPVNVLP 12
Db  374 FFFSPLINVLP 384

RESULT 3
ID  Q98886 PRELIMINARY; PRT; 144 AA.
AC  Q98886;
DT  01-OCT-2001 (TrEMBLrel. 18, Created)
DT  01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT  01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE  Hypothetical protein orf144.
GN  Name=orf144;
OS  Guillardia theta (Cryptomonas phi).
OG  Nucleomorph.
OC  Eukaryota; Cryptophyta; Cryptomonadaceae; Guillardia.
OX  NCBI_TaxID=55529;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=2123349; PubMed=11323671; DOI=10.1038/35074092;
RA  Douglas S., Zauner S., Fraunholz M., Beaton M., Penny S., Deng L.T.,
RA  Wu X., Reith M., Cavalier-Smith T., Maier U.G.;
RT  "The highly reduced genome of an enslaved algal nucleus.";
RL  Nature 410:1091-1096 (2001).
DR  EMBL; AF083031; AAK39696.1; -.
DR  PIR; D90125; D90125.
KW  Hypothetical protein; Nucleomorph.
SQ  SEQUENCE 144 AA; 17625 MW; 72649208661F4DA5 CRC64;

Query Match 72.6%; Score 45; DB 2; Length 144;
Best Local Similarity 66.7%; Pred. No. 7.5;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY  1 LFFFLPVNVLP 12
Db  24 IFFFLKKNILP 35

RESULT 4
ID  Q8UHQ7 PRELIMINARY; PRT; 62 AA.
AC  Q8UHQ7;
DT  01-JUN-2002 (TrEMBLrel. 21, Created)
DT  01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT  01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE  Hypothetical protein Atu0623.

```

```

GN  OrderedLocusNames=Atu0623;
OS  Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC  Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC  Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
OX  NCBI_TaxID=176229;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=Dupont;
RX  MEDLINE=21608550; PubMed=11743193; DOI=10.1126/science.1066804;
RA  Korda D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA  Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
RA  Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,
RA  Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
RA  Kutayavin T., Levy R., Li M.-J., McClelland E., Palmieri A.,
RA  Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
RA  Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA  Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA  Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA  Nester E.W.;
RT  "The genome of the natural genetic engineer Agrobacterium tumefaciens
RT  C58.";
RL  Science 294:2317-2323 (2001).
DR  EMBL; AE009030; AAL41639.1; -.
DR  PIR; A12652; A12652.
KW  Complete proteome; Hypothetical protein.
SQ  SEQUENCE 62 AA; 7135 MW; C67C9F18234FEAB6 CRC64;

Query Match 71.0%; Score 44; DB 2; Length 62;
Best Local Similarity 72.7%; Pred. No. 5.1;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY  2 PFPLPVNVLP 12
Db  5 FYSLPVNVLP 15

RESULT 5
ID  Q9G873 PRELIMINARY; PRT; 218 AA.
AC  Q9G873;
DT  01-MAR-2001 (TrEMBLrel. 16, Created)
DT  01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT  01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE  ABC transporter channel subunit.
GN  Name=yelV;
OS  Malawimonas jakobiformis.
OG  Mitochondrion.
OC  Eukaryota; Malawimonadidae; Malawimonas.
OX  NCBI_TaxID=136089;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=ATCC 50310;
RA  Burger G., O'Kelly C.J., Gray W.M.;
RL  Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR  EMBL; AF295546; AAG13700.1; -.
DR  GO; GO:0005739; C.mitochondrion; IEA.
KW  Mitochondrion.
SQ  SEQUENCE 218 AA; 25870 MW; EF7A162FEF8D674 CRC64;

Query Match 71.0%; Score 44; DB 2; Length 218;
Best Local Similarity 54.5%; Pred. No. 16;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY  1 LFFFLPVNVLP 11
Db  101 LFFFLPIITII 111

RESULT 6
ID  O24102 PRELIMINARY; PRT; 249 AA.
DT  01-JAN-1998 (TrEMBLrel. 05, Created)

```

DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Mtn4 protein (Fragment).
 GN Name=Mtn4;
 OS Medicago truncatula (Barrel medic).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.
 OX NCBI_TaxID=3880;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Root node;
 RX MEDLINE=96212994; PubMed=8634476;
 RA Gamas P.; de Carvalho Niebel F.; Lescure N.; Cullimore J.;
 RT "Use of a subtractive hybridization approach to identify new Medicago
 RT truncatula genes induced during root nodule development.";
 RL Mol. Plant Microbe Interact. 9:233-242(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Root node;
 RA Gamas P.;
 RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; Y15372; CAA75594.1; -.
 DR HSSP; P24337; IHYP.
 DR InterPro; IPR003612; AAI.
 DR Pfam; PF00234; Tryp_alpha_amy1; 1.
 DR SMART; SM00499; AAI; 1.
 FT NON TER 1
 FT SEQUENCE 249 AA; 26923 MW; 4BF9256A0FDD1318 CRC64;
 SQ SEQUENCE 249 AA; 26923 MW; 4BF9256A0FDD1318 CRC64;
 Query Match 71.0%; Score 44; DB 2; Length 249;
 Best Local Similarity 66.7%; Pred. No. 19;
 Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 Yy 1 LFFFLPVVVNLP 12
 Db 107 MFFFLPVVVPTP 118
 :||:|||||
 RESULT 7
 ID Q8FUV9 PRELIMINARY; PRT; 271 AA.
 AC Q8FUV9;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Polyamine ABC transporter, permease protein, putative.
 GN OrderedLocNames=BR11106;
 OS Brucella suis.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Brucellaceae; Brucella.
 OX NCBI_TaxID=29461;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1330 / Biovar 1;
 RX MEDLINE=22247741; PubMed=12271122; DOI=10.1073/pnas.192319099;
 RA Paulsen I.T.; Seshadri R.; Nelson K.E.; Eisen J.A.; Heidelberg J.F.;
 RA Read T.D.; Dodson R.J.; Unayam L.A.; Brinkac L.M.; Beanan M.J.;
 RA Daugherty S.C.; DeBoy R.T.; Durkin A.S.; Kolonay J.F.; Madupu R.;
 RA Nelson W.C.; Ayodeji B.; Kraul M.; Shetty J.; Malek J.A.;
 RA Van Aken S.E.; Riedmuller S.; Tettelin H.; Gill S.R.; White O.;
 RA Salzberg S.L.; Hoover D.L.; Lindler L.E.; Halling S.M.; Boyle S.M.;
 RA Fraser C.M.;
 RT "The Brucella suis genome reveals fundamental similarities between
 RT animal and plant pathogens and symbionts.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:13148-13153(2002).
 CC -!- FUNCTION: Part of a binding-protein-dependent transport system.
 CC Probably responsible for the translocation of the substrate across
 CC the membrane (By similarity).
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
 CC -!- SIMILARITY: Belongs to the binding-protein-dependent transport
 CC system permease family.
 EMBL; AE014602; AAN34268.1; -.

DR TIGR; BRA1106; -.
 DR GO; GO:0016020; C-membrane; IEA.
 DR GO; GO:0005215; F:transporter activity; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR000515; BPD_transp.
 DR Pfam; PF00528; BPD_transp_1; 1.
 DR PROSITE; PS50928; ABC_TM1; 1.
 KW Complete proteome; Transmembrane; Transport.
 SQ SEQUENCE 271 AA; 29759 MW; 6DDEC475BD4C1A83 CRC64;
 Query Match 71.0%; Score 44; DB 2; Length 271;
 Best Local Similarity 72.7%; Pred. No. 20;
 Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 Yy 1 LFFFLPVVVNL 11
 Db 14 LFFFLPVVVNL 24
 |||||:|:|:
 RESULT 8
 ID Q8VQL7 PRELIMINARY; PRT; 288 AA.
 AC Q8VQL7;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Putative ABC transporter permease protein B.
 GN Name=BaTn1953_ORF9;
 OS Brucella abortus.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Brucellaceae; Brucella.
 OX NCBI_TaxID=235;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=544;
 RA Bricker B.J.;
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Part of a binding-protein-dependent transport system.
 CC Probably responsible for the translocation of the substrate across
 CC the membrane (By similarity).
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
 CC -!- SIMILARITY: Belongs to the binding-protein-dependent transport
 CC system permease family.
 DR EMBL; AF454951; AAL59331.1; -.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0005215; F:transporter activity; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR Pfam; PF00528; BPD_transp_1; 1.
 DR PROSITE; PS50928; ABC_TM1; 1.
 KW Transmembrane; Transport.
 SQ SEQUENCE 288 AA; 31696 MW; B5C20EA208DCFD8E CRC64;
 Query Match 71.0%; Score 44; DB 2; Length 288;
 Best Local Similarity 72.7%; Pred. No. 21;
 Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 Yy 1 LFFFLPVVVNL 11
 Db 31 LFFFLPVVVNL 41
 |||||:|:|:
 RESULT 9
 ID Q9HIU9 PRELIMINARY; PRT; 525 AA.
 AC Q9HIU9;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Cytochrome b related protein.
 GN OrderedLocNames=Pal228;
 OS Thermoplasma acidophilum.
 OC Archaea; Euryarchaeota; Thermoplasmata; Thermoplasmatales;
 OC Thermoplasmataceae; Thermoplasma.

```

OX NCBI_TaxID=2303;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DSM 1728;
RX MEDLINE=20479972; PubMed=1029001; DOI=10.1038/35035069;
RA Ruepp A., Graml W., Santos-Martinez M.-L., Koretke K.K., Volker C.,
RA Meves H.-W., Frishman D., Stocker S., Lupas A.N., Baumeister W.;
RT "The genome sequence of the thermoacidophilic scavenger Thermoplasma
RT acidophilum.";
RL Nature 407:508-513 (2000).
DR EMBL; AL445086; CAC12332.1; -
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:006118; P:electron transport; IEA.
DR InterPro; IPR005797; Cytb_b6_N.
DR Pfam; PF00033; Cytochrom_B_N; 1.
KW Complete proteome.
SQ SEQUENCE 525 AA; 58544 MW; 145564FA78C665B7 CRC64;

Query Match 71.0%; Score 44; DB 2; Length 525;
Best Local Similarity 75.0%; Pred. No. 37;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 LFFFLPVVVVLP 12
Db 368 LFFFLPVVVVLP 379

RESULT 10
Q8A411 PRELIMINARY; PRT; 545 AA.
ID Q8A411
AC Q8A411;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative MFS transporter.
GN OrderedLocusNames=BT2793;
OS Bacteroides thetaiotaomicron.
OC Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;
OC Bacteroidaceae; Bacteroides.
OX NCBI_TaxID=818;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VPI-5482 / ATCC 29148;
RX MEDLINE=22550858; PubMed=12663928; DOI=10.1126/science.1080029;
RA Xu J., Bjursell M.K., Himrod J., Deng S., Carmichael L.K.,
RA Chiang H.C., Hooper L.V., Gordon J.I.;
RT "A genomic view of the human-Bacteroides thetaiotaomicron symbiosis.";
RL Science 299:2074-2076 (2003).
DR EMBL; AE016937; AAC077899.1; -
DR InterPro; IPR00412; ABC 2.
DR PROSITE; PS00890; ABC2_MEMBRANE; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 545 AA; 61484 MW; 976E2EF859984466 CRC64;

Query Match 67.7%; Score 42; DB 2; Length 545;
Best Local Similarity 70.0%; Pred. No. 85;
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LFFFLPVVVV 10
Db 24 LFFFLPVVLS 33

RESULT 11
Q7M9X9 PRELIMINARY; PRT; 585 AA.
ID Q7M9X9
AC Q7M9X9;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE SODIUM/SOLUTE SYMPORTER, PUTATIVE.
GN OrderedLocusNames=WS0601;

```

```

OS Wolinella succinogenes.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteraceae; Wolinella.
OX NCBI_TaxID=844;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DSMZ 1740;
RX MEDLINE=22882897; PubMed=14500908; DOI=10.1073/pnas.1932838100;
RA Baar C., Eppinger M., Raddatz G., Simon J., Lanz C., Klimmek O., Linke B.,
RA Nandakumar R., Gross R., Rosinus A., Keller H., Jagtap P., Linke B.,
RA Meyer F., Lederer H., Schuster S.C.;
RT "Complete genome sequence and analysis of Wolinella succinogenes.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:11690-11695 (2003).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -1- SIMILARITY: Belongs to the sodium:solute symporter (SSF) family.
DR EMBL; BX571658; CAE09732.1; -
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR001734; Na/solut_symport.
DR Pfam; PF00474; SSF, 1.
DR PROSITE; PS00457; NA_SOLUTE_SYMP_2; UNKNOWN_1.
DR PROSITE; PS0283; NA_SOLUTE_SYMP_3; 1.
KW Complete proteome; Transmembrane; Transport.
SQ SEQUENCE 585 AA; 63119 MW; 362293F8529FA91D CRC64;

Query Match 67.7%; Score 42; DB 2; Length 585;
Best Local Similarity 63.8%; Pred. No. 90;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 FFFFLPVVVVLP 12
Db 523 FFFFLPVVVVLP 533

RESULT 12
Q36231 PRELIMINARY; PRT; 66 AA.
ID Q36231
AC Q36231;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Partial putative ORF (Fragment).
OS Williopsis saturnus var. suaveolens.
OG Mitochondrion.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Williopsis.
OX NCBI_TaxID=58637;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CBS 255;
RA Rosek J.;
RL Submitted (JAN-1994) to the EMBL/GenBank/DBJ databases.
DR EMBL; X77238; CAA54455.1; -
DR PIR; S49114; S49114.
DR GO; GO:0005739; C:mitochondrion; IEA.
KW Mitochondrion.
FT NON_TER 1 1
SQ SEQUENCE 66 AA; 7586 MW; 78C9BCF9A31B94FB CRC64;

Query Match 66.1%; Score 41; DB 2; Length 66;
Best Local Similarity 50.0%; Pred. No. 18;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 1 LFFFLPVVVVLP 12
Db 37 LFFFLPVVVVLP 48

RESULT 13
Q96NK3 PRELIMINARY; PRT; 126 AA.
ID Q96NK3
AC Q96NK3;

```


01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein FLJ30690.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX PubMed=14702039; DOI=10.1038/ng1285;
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA Yamamoto J., Saito K., Kawai Y., Isono Y., Nakamura Y., Shiratori A.,
RA Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shiratori A.,
RA Sudo H., Hosiiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y.,
RA Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M., Yamazaki M.,
RA Ninomiya K., Ishibashi T., Yamashita H., Murakawa K., Fujimori K.,
RA Tanai H., Kimata M., Watanabe M., Hiraoa S., Chiba Y., Ishida S.,
RA Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T., Kusano J.,
RA Kanehori K., Takahashi-Fujii A., Hara H., Tanase T., Nomura Y.,
RA Togiya S., Komai F., Hara R., Takeuchi K., Arita M., Imose N.,
RA Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,
RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,
RA Moriya S., Momiya H., Satoh N., Takami S., Terashima Y., Suzuki O.,
RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
RA Fujimori Y., Koniyama M., Tashiro H., Tanigami A., Fujiwara T.,
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigetani K., Senba T.,
RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
RT "Complete sequencing and characterization of 21,243 full-length human
cDNAs.";
RL Nat. Genet. 36:40-45 (2004).
DR EMBL: AK055252; BAB70890.1; -;
SQ SEQUENCE 126 AA; 14391 MW; DIB23CDB0644847B CRC64;
Query Match 66.1%; Score 41; DB 2; Length 126;
Best Local Similarity 72.7%; Pred. No. 33;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
Qy 2 FFFLPVNVLP 12
Db 35 FFFLPVVSSLP 45
|||||:|
PRELIMINARY; PRT; 242 AA.
ID Q65223
AC Q65223;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Cyclic nucleotide-gated calmodulin-binding ion channel-like.
GN Name=OSJNBa0043B22.32;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Sasaki T., Matsumoto T., Katayose Y.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 6, BAC

RT clone:OSJNBa0043B22.";
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AP005470; BAD46124.1; -;
SQ SEQUENCE 242 AA; 27613 MW; 168B4C5E7991E868 CRC64;
Query Match 66.1%; Score 41; DB 2; Length 242;
Best Local Similarity 66.7%; Pred. No. 60;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
Qy 1 LFFFLPVNV 9
Db 74 LFFFLPVNV 82
|||||:|
PRELIMINARY; PRT; 376 AA.
ID Q8IM69
AC Q8IM69;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Exopolysphatase, putative.
GN ORFNames=PF14_0022;
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22255705; PubMed=12368864; DOI=10.1038/nature01097;
RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
RA Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA Chan M.S., Nene V., Shallow S.J., Suh B., Peterson J., Angiuoli S.,
RA Pertea M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,
RA Martin D.M., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA Fraser C.M., Barrell B.;
RT "Genome sequence of the human malaria parasite Plasmodium
falciparum.";
RL Nature 419:498-511 (2002).
DR EMBL: AE014816; AAN36634.1; -;
SQ SEQUENCE 376 AA; 45051 MW; 0FDE636678149F2C CRC64;
Query Match 66.1%; Score 41; DB 2; Length 376;
Best Local Similarity 60.0%; Pred. No. 90;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
Qy 1 LFFFLPVNV 10
Db 158 LFFFLPVNV 167
|||||:|
Search completed: April 12, 2005, 08:05:42
Job time : 59.9454 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 12, 2005, 07:49:31 ; Search time 84 Seconds
(without alignments)
69.064 Million cell updates/sec

Title: US-10-691-330-5

Perfect score: 81

Sequence: 1 DLEMPVLPPVPPFFV 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

- 1: Geneseqp1980s:*
- 2: Geneseqp1990s:*
- 3: Geneseqp2000s:*
- 4: Geneseqp2001s:*
- 5: Geneseqp2002s:*
- 6: Geneseqp2003as:*
- 7: Geneseqp2003bs:*
- 8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	81	100.0	15	4	AAB72504 Colostrin
2	81	100.0	15	4	AAB59322 Ewe colos
3	81	100.0	15	4	AAB72250 Colostrin
4	81	100.0	15	4	AAB72536 Colostrin
5	81	100.0	15	5	AAOI4581 Neural ce
6	81	100.0	15	5	AAAM51040 Colostrin
7	81	100.0	15	5	AAE20232 Colostrin
8	81	100.0	15	8	ADN60299 Constitue
9	81	100.0	15	8	ADS74398 Ovine col
10	81	100.0	16	4	AAB59352 Ewe colos
11	57	70.4	10	4	AAE07187 Colostrin
12	51	63.0	10	4	AAE07197 Modified
13	48	59.3	232	7	ADB64957 Human pro
14	46	56.8	180	7	ADM26588 Hyperther
15	45	55.6	55	4	AAU60386 Propionib
16	45	55.6	55	6	ABM56905 Propionib
17	45	55.6	89	3	AAAG11334 Arabidops
18	45	55.6	91	3	AAAG11333 Arabidops
19	45	55.6	103	3	AAAG11332 Arabidops
20	45	55.6	446	6	ABU49258 Protein e
21	45	55.6	619	5	ABP69657 Human pol
22	45	55.6	765	5	ADF74168 Human nov
23	45	55.6	803	7	ABO81733 Pseudomon
24	45	55.6	902	5	ABP65500 Bifidobac
25	44	54.3	247	4	ABG13624 Novel hum

26	44	54.3	259	6	ADA33586	Ada33586 Acinetoba
27	44	54.3	388	8	ADQ08702	Adq08702 Ciona int
28	43.5	53.7	508	8	ADR86525	Adr86525 Aspergill
29	42	51.9	54	4	ABG06766	Abg06766 Novel hum
30	42	51.9	99	8	ADP29919	Adp29919 Human sec
31	42	51.9	141	6	ABP79310	Abp79310 N. gonorr
32	42	51.9	232	4	ABB60555	Abb60555 Drosophil
33	42	51.9	242	7	ABO79615	AbO79615 Pseudomon
34	42	51.9	457	8	ADS29748	Ads29748 Bacterial
35	42	51.9	469	4	ABB61392	Abb61392 Drosophil
36	41.5	51.2	78	3	AAAG01794	Aag01794 Human sec
37	41.5	51.2	86	3	AAAY86518	Aay86518 Human gen
38	41.5	51.2	86	6	ABO53682	ABO53682 Novel hum
39	41.5	51.2	119	4	AAAG91178	Aam91178 Human imm
40	41	50.6	63	4	ABG14725	Abg14725 Novel hum
41	41	50.6	83	4	ABG10974	Abg10974 Novel hum
42	41	50.6	101	4	ABB11472	Abb11472 Human lam
43	41	50.6	133	3	AAG48755	Aag48755 Arabidops
44	41	50.6	153	4	ABB71049	Abb71049 Drosophil
45	41	50.6	159	3	AAG48761	Aag48761 Arabidops

ALIGNMENTS

RESULT 1
AAB72504
ID AAB72504 standard; peptide; 15 AA.
XX AAB72504;
AC AAB72504;
XX 09-MAY-2001 (first entry)
DT Colostrin peptide #5.
DE Colostrin peptide #5.
XX Dermatological; oxidative stress regulator; colostrin.
KW Unidentified.
XX OS
XX WO200112650-A2.
XX 22-FEB-2001.
XX 17-AUG-2000; 2000WO-US022665.
XX 17-AUG-1999; 99US-0149310P.
XX (TEXA) UNIV TEXAS SYSTEM.
XX Stanton GJ, Hughes TK, Boldogh I;
XX WPI; 2001-218342/22.
XX Modulating oxidative stress level in a cell, involves contacting the cell with an oxidative stress regulator selected from colostrin, its constituent peptide, analog or their combinations.
XX Claim 6; Page 25; 48pp; English.

The present invention relates to a method for modulating the oxidative stress level in a cell or a patient, comprising contacting the cell with, or administering to the patient, an oxidative stress regulator selected from colostrin, or its constituent peptide (e.g. the present peptide), to change the level of an oxidising species in the cell. The method can be used to treat oxidative damage to skin, by decreasing or preventing an increase in the level of damage to a biomolecule of the patient

Sequence 15 AA;

Query Match 100.0%; Score 81; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. NO. 1.4e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

PA (TEXA) UNIV TEXAS SYSTEM.
 XX Boldogh I;
 XX WPI; 2001-226545/23.
 DR Use of colostrinin, its constituent peptide or analog as a neural cell
 PT regulator, for promoting neural cell differentiation and treating damaged
 PT neural cells in a patient.
 XX
 XX Claim 6; Page 21; 35pp; English.
 PS
 CC The present invention relates to a method for promoting neural cell
 CC differentiation and treating damaged neural cells, using colostrinin and
 CC colostrinin constituent peptides (e.g. the present peptide) as a neural
 CC cell regulator. Colostrinin is a polypeptide complex found in colostrum
 XX
 XX Sequence 15 AA;
 SQ

Query Match 100.0%; Score 81; DB 4; Length 15;
 Best Local Similarity 100.0%; Pred. No. 1.4e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DLEMPVLPVEPPFPV 15
 |||||
 Db 1 DLEMPVLPVEPPFPV 15
 |||||

RESULT 5
 AA014581
 ID AA014581 standard; peptide; 15 AA.
 AC AA014581;
 XX
 XX 27-MAY-2002 (first entry)
 DT
 XX
 XX Neural cell regulatory colostrinin peptide 5.
 DE
 XX
 XX Neural cell differentiation; neural cell regulator; colostrinin peptide;
 KW neural cell formation; proline-rich polypeptide aggregate; colostrum;
 KW neural cell treatment.
 XX
 XX Unidentified.
 OS
 XX
 XX Key Location/Qualifiers
 FH Modified-site 15 /note= "Optional C-terminal amide"
 FT
 XX WO200213851-A1.
 PN
 XX 21-FEB-2002.
 PD
 XX 17-AUG-2000; 2000WO-US022777.
 XX
 XX 17-AUG-2000; 2000WO-US022777.
 PR
 XX (TEXA) UNIV TEXAS SYSTEM.
 PA Boldogh I, Stanton JG, Hughes TK;
 PA WPI; 2002-269152/31.
 DR
 XX Promoting cell differentiation in a patient involves use of blood cell
 PT regulator selected from colostrinin, its constituent peptide and/or
 PT analog.
 XX
 XX Claim 7; Page 21; 37pp; English.
 PS
 CC The invention comprises a method for promoting cell differentiation (e.g.
 CC neural cell differentiation). The method involves contacting cells with a
 CC neural cell regulator (i.e. a colostrinin peptide) in order to change the
 CC cells in morphology to form neural cells. Colostrinin is a proline-rich
 CC polypeptide aggregate that is present in colostrum. The method of the

CC invention is useful for promoting the differentiation of cells and for
 CC treating damaged neural cells in a patient. The present amino acid
 CC sequence represents a specifically claimed colostrinin peptide used in
 CC the method of the invention
 XX
 XX Sequence 15 AA;
 SQ

Query Match 100.0%; Score 81; DB 5; Length 15;
 Best Local Similarity 100.0%; Pred. No. 1.4e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DLEMPVLPVEPPFPV 15
 |||||
 Db 1 DLEMPVLPVEPPFPV 15
 |||||

RESULT 6
 AAM51040
 ID AAM51040 standard; peptide; 15 AA.
 AC AAM51040;
 XX
 XX 30-MAY-2002 (first entry)
 DT
 XX
 XX Colostrinin constituent peptide.
 DE
 XX
 XX Colostrinin; colostrum; immunomodulator; cardiovascular;
 KW blood cell regulator; cytokine inducer; beta-casein; human.
 KW
 XX
 XX Homo sapiens.
 OS
 XX
 XX Key Location/Qualifiers
 FH Modified-site 15 /note= "optional C-terminal amidation"
 FT
 XX WO200213849-A1.
 PN
 XX 21-FEB-2002.
 PD
 XX 17-AUG-2000; 2000WO-US022775.
 XX
 XX 17-AUG-2000; 2000WO-US022775.
 PR
 XX (TEXA) UNIV TEXAS SYSTEM.
 PA (REG-) REGEN THERAPEUTICS PLC.
 PA
 XX Stanton GJ, Hughes TK, Boldogh I, Georgiades J;
 FI WPI; 2002-269150/31.
 XX
 XX Modulation of blood cell proliferation in a patient involves use of blood
 PT cell regulator selected from colostrinin, its constituent peptide and/or
 PT analog.
 XX
 XX Claim 1; Page 34; 54pp; English.
 PS
 CC The present sequence is that of a colostrinin constituent peptide that is
 CC preferred for use as an immunological regulator and as a blood cell
 CC regulator in claimed methods of the invention. It is classified as having
 CC a beta-casein homologue precursor. Methods are claimed for: inducing a
 CC cytokine in a cell by contact with an immunological regulator, where the
 CC cell is present in a cell culture, a tissue, an organ or an organism, and
 CC the cell is mammalian, including human; modulating an immune response in
 CC a cell by contact with the immunological regulator under conditions
 CC effective to induce a cytokine; modulating an immune response in a
 CC patient by administering an immunological regulator under conditions
 CC effective to induce a cytokine, where the immunological regulator is
 CC administered topically or as part of a dietary supplement, and where the
 CC immune response is specific or non specific, an interferon response or an
 CC antibody response; modulating blood cell proliferation by contacting
 CC blood cells with a blood cell regulator, where the blood cells are
 CC present in a cell culture or an organism, are mammalian or human, and
 CC where the blood cells are increased in number or differentiated; and a

CC method for modulating blood cell proliferation in a patent. A claimed
 CC cytokine-inducing composition comprises a pharmaceutical carrier and an
 CC active agent such as the present peptide. Cytokines induced by this
 CC peptide in human leucocyte cultures include interferon-gamma, tumour
 CC necrosis factor-alpha, interleukin-6 and interleukin-10
 XX
 XX
 SQ Sequence 15 AA;

Query Match 100.0%; Score 81; DB 5; Length 15;
 Best Local Similarity 100.0%; Pred. No. 1.4e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DLEMPVLPEPPFV 15
 |||||
 Db 1 DLEMPVLPEPPFV 15
 |||||

RESULT 7
 ID AAE20232 standard; peptide; 15 AA.
 XX
 AC AAE20232;
 XX
 DT 18-JUN-2002 (first entry)
 XX
 DE Colostrin constituent peptide #5.
 XX
 KW Blood cell regulator; colostrin; constituent peptide; oxidative stress;
 KW therapy; oxidative damage; skin; aging; wound healing; cell replacement;
 KW tissue; organ; cosmetic procedure; repair; regeneration; preservation;
 KW transplantation; implantation; dermatological; vulnery.
 XX
 OS Unidentified.
 XX

Key Location/Qualifiers
 FH 15
 FT Modified-site /note= "Optionally C-terminal amide"
 FT 15
 PN WO200213850-A1.
 XX
 XX 21-FEB-2002.
 XX
 XX 17-AUG-2000; 2000WO-US022776.
 XX
 XX 17-AUG-2000; 2000WO-US022776.
 XX
 XX (TEXA) UNIV TEXAS SYSTEM.
 XX
 XX Stanton GJ, Hughes TK, Boldogh I;
 XX WPI; 2002-269151/31.
 XX
 XX Composition useful for the modulation of blood cell proliferation in a
 PT patient comprises a blood cell regulator selected from colostrin, its
 PT constituent peptide and/or analog.
 XX
 XX Claim 6; Page 25; 51pp; English.
 PS
 XX

The invention relates to a composition which comprises a blood cell
 CC regulator selected from colostrin, its constituent peptide and/or
 CC analogue. The invention is used for modulating the oxidative stress level
 CC in a cell e.g. mammalian or human cell present in a cell culture, tissue,
 CC organ, or organism; or for treating oxidative damage to the skin of a
 CC patient e.g. animal or human; to modulate oxidative stress during/ after
 CC a premature birth or normal birth, preventing/delaying aging in a
 CC patient, enhancing wound healing, and the reduction of side effects of
 CC cosmetic procedures. The method changes the level of an oxidising species
 CC in the cell, such as decreases or prevents increase in the level of
 CC damage to a biomolecule of the patient selected from DNA, protein and/or
 CC lipid compared to the same conditions when the oxidative stress
 CC regulator is not present. The modulation of oxidative stress results in
 CC enhanced repair, regeneration, and replacement of cells, tissues and
 CC organs (e.g. kidney, liver, pancreas, skin, and the other internal and

CC external organs), as well as enhanced preservation of such organs for
 CC transplantation, implantation, or scientific research. The present
 CC sequence is a colostrin constituent peptide
 XX
 XX Sequence 15 AA;

Query Match 100.0%; Score 81; DB 5; Length 15;
 Best Local Similarity 100.0%; Pred. No. 1.4e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DLEMPVLPEPPFV 15
 |||||
 Db 1 DLEMPVLPEPPFV 15
 |||||

RESULT 8
 ID ADN60299 standard; peptide; 15 AA.
 XX
 AC ADN60299;
 XX
 DT 29-JUL-2004 (first entry)
 XX
 DE Constituent peptide of colostrin SEQ ID NO:5.
 XX
 KW modulator; colostrin; intracellular signaling-molecule modulator;
 KW 4HNE-mediated lipid peroxidation down regulation; apoptosis inhibition;
 KW DNA damage; beta-amyloid; retinoic acid; cytostatic; 4HNE inhibitor;
 KW 4HNE-protein adduct formation reduction;
 KW 4HNE-mediated glutathione depletion inhibition; p53 inhibition;
 KW C-Jun NH2-terminal kinase inhibition.
 XX
 OS Synthetic.
 XX
 XX WO2004037851-A2.
 XX
 XX 06-MAY-2004.
 XX
 XX 22-OCT-2003; 2003WO-US033423.
 XX
 XX 22-OCT-2002; 2002US-0420369P.
 PR
 XX (TEXA) UNIV TEXAS SYSTEM.
 PA (BOLD/) BOLDOGH I.
 PA (STAN/) STANTON J G.
 PA (GEOR/) GEORGIADIS J A.
 PA (HUGH/) HUGHES T K.
 PA (KRUIZ/) KRUIZEL M.
 XX
 XX Boldogh I, Stanton JG, Georgiades JA, Hughes TK, Kruzel M;
 XX WPI; 2004-365494/34.
 DR
 XX Use of colostrin for e.g. modulating an intracellular signaling
 PT molecule in a cell, down regulating 4HNE-mediated lipid peroxidation in a
 PT cell, inhibiting apoptosis in a cell, or protecting against DNA damage in
 PT a cell.
 XX
 XX Claim 6; SEQ ID NO 5; 46pp; English.
 PS
 XX The present invention describes the use of a modulator selected from
 CC colostrin, its constituent peptide, its active analogue, and a
 CC combination of these, for modulating an intracellular signaling molecule
 CC in a cell, down regulating 4HNE-mediated lipid peroxidation in a cell,
 CC inhibiting apoptosis in a cell, protecting against DNA damage in a cell,
 CC and reducing the toxic effect of beta-amyloid or retinoic acid on a cell.
 CC The modulator has cytostatic activity, and can be used as a 4HNE
 CC inhibitor. The modulator is useful in the manufacture of a medicament for
 CC reducing 4HNE-protein adduct formation, inhibiting 4HNE-mediated
 CC glutathione depletion, inhibiting 4HNE-induced activation of p53 protein,
 CC and/or inhibiting 4HNE-induced activation of C-Jun NH2-terminal kinases.
 CC Colostrin, or its constituent peptide or active analogue is useful for
 CC inhibiting apoptosis in a cell, protecting against DNA damage in a cell,

CC or reducing the toxic effect of beta-amyloid or retinoic acid on a cell.
 CC The present sequence represents a synthetic constituent peptide of
 CC colostrinin, which can be used as a modulator in the present invention.

XX SQ Sequence 15 AA;
 Query Match 100.0%; Score 81; DB 8; Length 15;
 Best Local Similarity 100.0%; Pred. No. 1.4e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DLEMPVLVPEPPFV 15
 |||||
 Db 1 DLEMPVLVPEPPFV 15

RESULT 9
 ADS74398
 ID ADS74398 standard; peptide; 15 AA.

XX AC ADS74398;
 XX DT 16-DEC-2004 (first entry)
 XX KW Ovine colostrinin peptide.
 XX DE Colostrum; colostrinin; sheep; peptide purification.
 XX OS Ovis aries.

XX PN WO2004081038-A1.

XX PD 23-SEP-2004.

XX PF 10-MAR-2004; 2004WO-GB001014.

XX PR 11-MAR-2003; 2003GB-00005552.

XX PA 08-MAR-2004; 2004GB-00005190.

XX PI (REG-) REGEN THERAPEUTICS PLC.

XX PS Georgiades JA, Polanowski A, Wilusz T, Kruzel ML;

XX PT WPI; 2004-677519/66.

XX PT Recovering peptides such as colostrinin from mammalian colostrum, by

XX PT mixing colostrum with alcohol to form alcohol phase containing peptides

XX PT and precipitate, separating alcohol phase from precipitate, and

XX PT recovering alcohol phase.

XX PS Example; SEQ ID NO 15; 41pp; English.

XX CC The present sequence is that of a peptide that can be recovered from

XX CC ovine colostrinin using the method of the invention. The invention

XX CC provides a method for the recovery of peptides (especially colostrinin)

XX CC from colostrum in substantially pure, biologically active form and in

XX CC high yield. The method involves mixing the colostrum with an alcohol to

XX CC form an alcohol phase containing the colostrinin and a precipitate

XX CC containing higher molecular weight caseins and other proteins. Best

XX CC results are obtained using methanol or ethanol of at least 80%, and

XX CC preferably up to 100%, purity. The alcohol phase is then separated from

XX CC the precipitate, and the colostrinin is separated from the alcohol,

XX CC preferably by evaporation, to form a colostrinin-rich phase, which is

XX CC recovered. A precipitation agent, such as ammonium sulfate, may be added

XX CC either to the alcohol phase or, preferably, to the colostrinin-rich phase

XX CC to induce precipitation of the colostrinin peptides. The method is

XX CC generally applicable to the separation of peptides from fluids containing

XX CC higher molecular weight proteins, lipids, carbohydrates and/or nucleic

XX CC acids. In an example from the invention, the antigenic profile of

XX CC peptides recovered from sheep colostrum using the alcohol precipitation

XX CC methods was determined by ELISA using antibodies prepared against 9

XX CC synthetic peptides, including a peptide having the present sequence

XX CC (denoted antigen class B-8).

SQ Sequence 15 AA;

Query Match 100.0%; Score 81; DB 8; Length 15;
 Best Local Similarity 100.0%; Pred. No. 1.4e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DLEMPVLVPEPPFV 15
 |||||
 Db 1 DLEMPVLVPEPPFV 15

RESULT 10

AAB59352

ID AAB59352 standard; peptide; 16 AA.

XX AC AAB59352;

XX DT 21-MAR-2001 (first entry)

XX DE Ewe colostrinin peptide fragment derived sequence #12.

XX KW Sheep; colostrinin; proline rich polypeptide; colostrum; immune disorder;

XX KW central nervous system disorder; dietary supplement; beta-amyloid plaque.

XX OS Ovis sp.

XX PN WO2000075173-A2.

XX PD 14-DEC-2000.

XX PF 02-JUN-2000; 2000WO-GB002128.

XX PR 02-JUN-1999; 99GB-00012852.

XX PA (REG-) REGEN THERAPEUTICS PLC.

XX PI Georgiades JA;

XX PS WPI; 2001-071058/08.

XX PT Peptides having an N-terminal amino acid sequence isolated from

XX PT colostrinin for treating e.g. disorders of the central nervous system and

XX PT immune system, viral and bacterial infections, and diseases characterized

XX PT by amyloid plaques.

XX PS Claim 8; Page 27; 63pp; English.

XX CC The present invention provides the sequences of a number of peptides

XX CC found in ewe's colostrinin. Colostrinin is the proline-rich polypeptide

XX CC fragment of colostrum. These peptides can be used in the treatment of

XX CC central nervous system disorders such as senile dementia, Parkinson's

XX CC disease, Alzheimer's disease, psychosis and neurosis, immune system

XX CC disorders such as bacterial and viral infections, to improve the

XX CC development of a child's immune system, as a dietary supplement, and to

XX CC promote the dissolution of beta-amyloid plaques

XX SQ Sequence 16 AA;

Query Match 100.0%; Score 81; DB 4; Length 16;

Best Local Similarity 100.0%; Pred. No. 1.5e-05;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DLEMPVLVPEPPFV 15

2 DLEMPVLVPEPPFV 16

Db

RESULT 11

AAE07187

ID AAE07187 standard; peptide; 10 AA.

XX AC AAE07187;

XX XX

06-NOV-2001 (first entry)
Colostrinin peptide 3.
Colostrinin; nootropic; neuroprotective; immunomodulatory; antibacterial;
Parkinson's disease; Alzheimer's disease; mental disorder; food additive;
central nervous system disorder; neurodegenerative disorder; weight loss;
beta-amyloid plaque; psychosis; neurosis; cachexia; motor neuron disease;
acquired immunological deficiency; neurological disorder; dementia;
antiviral; cyclic.
Synthetic.
Key Modified-site 1 Location/Qualifiers
/note= "N-terminal acetyl; this residue forms a cyclic
linkage with Pro found at the C-terminal end"
WO200155199-A1.
02-AUG-2001.
26-JAN-2001; 2001WO-GB000329.
26-JAN-2000; 2000GB-00001825.
(REG-) REGEN THERAPEUTICS PLC.
Georgiades JA;
WPI; 2001-488775/53.
Peptide useful as an interalia in the treatment of e.g. disorders of the
immune system and the central nervous system comprises ten amino-terminal
amino acid sequence derived from peptides present in colostrinin.
Claim 1; Page 15; 40pp; English.
The invention relates to colostrinin peptide fragments which are useful,
inter alia, in the treatment of chronic disorders of the immune system
and the central nervous system. Colostrinin peptides are used as a
medicament in the treatment of neurological disorders e.g., dementia,
neurodegenerative disorders e.g., Alzheimer's disease, motor neuron
disease e.g., Parkinson's disease, mental disorders e.g. psychosis and
neurosis, in acquired immunological deficiencies, chronic bacterial and
viral infections and diseases characterised by the presence of beta-
amyloid plaques and as a dietary supplement for babies, small children,
adults and senile persons, who have been subjected to chemotherapy or
have suffered from cachexia or weight loss due to the chronic disease.
Colostrinin peptides are also used as food additives and as an auxiliary
withdrawal treatment for drug addicts, after a period of detoxification
and in persons dependent on stimulants. Colostrinin peptides are used to
prepare antibodies and to treat emotional disturbances, e.g. emotional
disturbances of psychiatric patients in a state of depression. These
colostrinin peptides improves the development of immune system in a new
born child and to correct the immunological deficiencies in a child. The
present sequence is colostrinin peptide 3 related to the invention
Sequence 10 AA;
Query Match 70.4%; Score 57; DB 4; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.041;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 5 PVLVPEPPFF 14
| | | | | | | |
Db 1 PVLVPEPPFF 10
RESULT 12
AAE07197
ID AAE07197 standard; peptide; 10 AA.
AC AAE07197;
XX
XX
XX 06-NOV-2001 (first entry)
XX
XX Modified colostrinin cyclic peptide #3.
Colostrinin; nootropic; neuroprotective; immunomodulatory; antibacterial;
Parkinson's disease; Alzheimer's disease; mental disorder; food additive;
central nervous system disorder; neurodegenerative disorder; weight loss;
beta-amyloid plaque; psychosis; neurosis; cachexia; motor neuron disease;
acquired immunological deficiency; neurological disorder; dementia;
antiviral.
Unidentified.
Key Modified-site 1 Location/Qualifiers
/note= "N-terminal acetyl; this residue forms a cyclic
linkage with Pro found at the C-terminal end"
WO200155199-A1.
02-AUG-2001.
26-JAN-2001; 2001WO-GB000329.
26-JAN-2000; 2000GB-00001825.
(REG-) REGEN THERAPEUTICS PLC.
Georgiades JA;
WPI; 2001-488775/53.
Peptide useful as an interalia in the treatment of e.g. disorders of the
immune system and the central nervous system comprises ten amino-terminal
amino acid sequence derived from peptides present in colostrinin.
Claim 1; Page 15; 40pp; English.
The invention relates to colostrinin peptide fragments which are useful,
inter alia, in the treatment of chronic disorders of the immune system
and the central nervous system. Colostrinin peptides are used as a
medicament in the treatment of neurological disorders e.g., dementia,
neurodegenerative disorders e.g., Alzheimer's disease, motor neuron
disease e.g., Parkinson's disease, mental disorders e.g. psychosis and
neurosis, in acquired immunological deficiencies, chronic bacterial and
viral infections and diseases characterised by the presence of beta-
amyloid plaques and as a dietary supplement for babies, small children,
adults and senile persons, who have been subjected to chemotherapy or
have suffered from cachexia or weight loss due to the chronic disease.
Colostrinin peptides are also used as food additives and as an auxiliary
withdrawal treatment for drug addicts, after a period of detoxification
and in persons dependent on stimulants. Colostrinin peptides are used to
prepare antibodies and to treat emotional disturbances, e.g. emotional
disturbances of psychiatric patients in a state of depression. These
colostrinin peptides improves the development of immune system in a new
born child and to correct the immunological deficiencies in a child. The
present sequence is colostrinin peptide 3 related to the invention
Sequence 10 AA;
Query Match 70.4%; Score 57; DB 4; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.041;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 5 PVLVPEPPFF 14
| | | | | | | |
Db 1 PVLVPEPPFF 10
RESULT 12
AAE07197
ID AAE07197 standard; peptide; 10 AA.
AC AAE07197;
XX
XX
XX 06-NOV-2001 (first entry)
XX
XX Modified colostrinin cyclic peptide #3.
Colostrinin; nootropic; neuroprotective; immunomodulatory; antibacterial;
Parkinson's disease; Alzheimer's disease; mental disorder; food additive;
central nervous system disorder; neurodegenerative disorder; weight loss;
beta-amyloid plaque; psychosis; neurosis; cachexia; motor neuron disease;
acquired immunological deficiency; neurological disorder; dementia;
antiviral; cyclic.
Synthetic.
Key Modified-site 1 Location/Qualifiers
/note= "N-terminal acetyl; this residue forms a cyclic
linkage with Pro found at the C-terminal end"
WO200155199-A1.
02-AUG-2001.
26-JAN-2001; 2001WO-GB000329.
26-JAN-2000; 2000GB-00001825.
(REG-) REGEN THERAPEUTICS PLC.
Georgiades JA;
WPI; 2001-488775/53.
Peptide useful as an interalia in the treatment of e.g. disorders of the
immune system and the central nervous system comprises ten amino-terminal
amino acid sequence derived from peptides present in colostrinin.
Claim 1; Page 15; 40pp; English.
The invention relates to colostrinin peptide fragments which are useful,
inter alia, in the treatment of chronic disorders of the immune system
and the central nervous system. Colostrinin peptides are used as a
medicament in the treatment of neurological disorders e.g., dementia,
neurodegenerative disorders e.g., Alzheimer's disease, motor neuron
disease e.g., Parkinson's disease, mental disorders e.g. psychosis and
neurosis, in acquired immunological deficiencies, chronic bacterial and
viral infections and diseases characterised by the presence of beta-
amyloid plaques and as a dietary supplement for babies, small children,
adults and senile persons, who have been subjected to chemotherapy or
have suffered from cachexia or weight loss due to the chronic disease.
Colostrinin peptides are also used as food additives and as an auxiliary
withdrawal treatment for drug addicts, after a period of detoxification
and in persons dependent on stimulants. Colostrinin peptides are used to
prepare antibodies and to treat emotional disturbances, e.g. emotional
disturbances of psychiatric patients in a state of depression. These
colostrinin peptides improves the development of immune system in a new
born child and to correct the immunological deficiencies in a child. The
present sequence is colostrinin peptide 3 related to the invention
Sequence 10 AA;
Query Match 63.0%; Score 51; DB 4; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.33;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 5 PVLVPEPPFF 13
| | | | | | | |
Db 2 PVLVPEPPFF 10
RESULT 13
ADB64957
ID ADB64957 standard; protein; 232 AA.
XX
XX ADB64957;
AC
XX

DT 04-DEC-2003 (first entry)
 XX Human protein encoded by clone PROST20054660.
 DE
 XX Human; pharmaceutical; diagnostic; gene therapy; tissue regeneration;
 KW cell regeneration; membrane protein; signal transduction-related protein;
 KW transcription-related protein; osteoporosis; neurological disease;
 KW cancer; tumour.
 XX
 XX Homo sapiens.
 OS
 XX EP1308459-A2.
 PN
 XX 07-MAY-2003.
 PD
 XX 28-MAR-2002; 2002EP-00007401.
 PF
 XX 05-NOV-2001; 2001JP-00379298.
 PR
 XX 25-JAN-2002; 2002US-00350978.
 PP
 XX (HELI-) HELIX RES INST.
 PA (REAS-) RES ASSOC BIOTECHNOLOGY.
 XX
 XX Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
 PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
 PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;
 XX
 XX WPI; 2003-450961/43.
 DR N-PSDB; ADB62987.
 XX
 XX New polynucleotides and polypeptides, useful for developing a diagnostic
 PT marker or medicines for regulation of their expression and activity, or
 PT as targets of gene therapy.
 XX
 XX Claim 1; Page; 222pp; English.
 PS
 XX The invention discloses a polynucleotide comprising a sequence selected
 CC from 1970 fully defined nucleotide sequences which encode novel
 CC polypeptides. Also claimed is a polypeptide encoded by the polynucleotide
 CC or its partial peptide, an antibody binding to the polypeptide or peptide
 CC of the polynucleotide, immunologically assaying the polypeptide or
 CC peptide of the polynucleotide by contacting the polypeptide or peptide
 CC with the antibody of the encoded protein, and observing the binding
 CC between the two, a transformant carrying the polynucleotide in an
 CC expressible manner and an antisense polynucleotide. The oligonucleotide
 CC is useful as a primer for synthesising the polynucleotide, or as a probe
 CC for detecting the polynucleotide. The polynucleotides and encoded
 CC proteins are useful as pharmaceutical agents and many disease-related
 CC genes may be included in them, for developing a diagnostic marker or
 CC medicines for regulation of their expression and activity, or as targets
 CC of gene therapy. The genes are involved in tissue and/or cell
 CC regeneration. Membrane proteins, signal transduction-related proteins,
 CC transcription-related proteins, disease-related proteins and genes
 CC encoding them can be used as indicators for diseases (e.g. osteoporosis,
 CC neurological diseases, cancer, tumours). The cDNA may be used to regulate
 CC the activity or expression of the encoded protein to treat diseases. The
 CC sequence presented is a protein of the invention. Note: Some of the
 CC sequence data for this patent is not represented in the printed
 CC specification, but is based on sequence information supplied by the
 CC European Patent Office.
 XX
 XX Sequence 232 AA;
 SQ
 Query Match 59.3%; Score 48; DB 7; Length 232;
 Best Local Similarity 66.7%; Pred. NO. 28;
 Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 Qy 3 EMPVLVPEPPF 14
 Db 209 KFPVLVPHWPPF 220
 RESULT 14

ADM26588
 ID ADM26588 standard; protein; 180 AA.
 XX
 AC ADM26588;
 XX
 DT 20-MAY-2004 (first entry)
 XX
 DE Hyperthermophile Methanopyrus kandleri protein #1194.
 XX
 KW hyperthermophile; protein stability enhancement;
 KW protein activity enhancement.
 XX
 OS Methanopyrus kandleri.
 PN
 XX WO2003076575-A2.
 PD
 XX 18-SEP-2003.
 PF
 XX 04-MAR-2003; 2003WO-US006664.
 PR
 XX 04-MAR-2002; 2002US-0361742P.
 PR
 XX 14-MAY-2002; 2002US-0380423P.
 PR
 XX 16-SEP-2002; 2002US-0410974P.
 XX
 XX (FIDE-) FIDELITY SYSTEMS INC.
 PA (MALY/) MALYKH A.
 XX
 XX Slesarev AI, Pavlov A, Pavlova N, Kozyavkin S;
 PI
 XX WPI; 2003-748383/70.
 DR N-PSDB; ADM27081.
 XX
 XX New isolated nucleic acids encoding any of about 1700 Methanopyrus
 PT kandleri proteins, and the encoded proteins, useful as a medicaments or
 PT as diagnostic agents.
 XX
 XX Claim 31; SEQ ID NO 1194; 1023pp; English.
 PS
 XX The invention comprises the amino acid sequence of proteins from the
 CC hyperthermophile Methanopyrus kandleri, the invention also comprises the
 CC complete genome from Methanopyrus kandleri. The Methanopyrus kandleri
 CC proteins of the invention are useful for enhancing the stability and/or
 CC activity of other proteins. The Methanopyrus kandleri genome is useful in
 CC a variety of diagnostic and analytical methods. The present amino acid
 CC sequence represents a Methanopyrus kandleri protein of the invention.
 XX
 XX Sequence 180 AA;
 SQ
 Query Match 56.8%; Score 46; DB 7; Length 180;
 Best Local Similarity 71.4%; Pred. No. 42;
 Matches 10; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 Qy 2 LEMPVLVPEPPFV 15
 Db 29 LECSVLVPPPEPFV 42
 RESULT 15
 AAU60386
 ID AAU60386 standard; protein; 55 AA.
 XX
 AC AAU60386;
 XX
 DT 27-FEB-2002 (first entry)
 XX
 DE Propionibacterium acnes immunogenic protein #21282.
 XX
 KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteopathic; neuroprotectant.
 XX
 OS Propionibacterium acnes.

```

XX WO200181581-A2.
XX
XX
XX PD 01-NOV-2001.
XX
XX PF 20-APR-2001; 2001WO-US012865.
XX
XX 21-APR-2000; 2000US-0199047P.
XX PR 02-JUN-2000; 2000US-0208841P.
XX PR 07-JUL-2000; 2000US-0216747P.
XX
XX PA (CORI-) CORIXA CORP.
XX
XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
XX PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
XX DR WPI; 2001-616774/71.
XX DR N-PSDB; AAS59609.
XX
XX Propionibacterium acnes polypeptides and nucleic acids useful for
XX PT vaccinating against and diagnosing infections, especially useful for
XX PT treating acne vulgaris.
XX
XX Example 1; SEQ ID NO 21581; 1069pp; English.
XX
XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
XX polypeptides. The proteins and their associated DNA sequences are used in
XX the treatment, prevention and diagnosis of medical conditions caused by
XX P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
XX pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
XX P. acnes is also involved in infections of bone, joints and the central
XX nervous system, however it is particularly involved in the inflammatory
XX lesions associated with acne vulgaris. A method for detecting the
XX presence or absence of P. acnes in a patient comprises contacting a
XX sample with a binding agent that binds to the proteins of the invention
XX and determining the amount of bound protein in the sample. The
XX polypeptides may be used as antigens in the production of antibodies
XX specific for P. acnes proteins. These antibodies can be used to
XX downregulate expression and activity of P. acnes polypeptides and
XX therefore treat P. acnes infections. The antibodies may also be used as
XX diagnostic agents for determining P. acnes presence, for example, by
XX enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
XX this patent did not form part of the printed specification, but was
XX obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX SQ Sequence 55 AA;
XX
XX Query Match 55.6%; Score 45; DB 4; Length 55;
XX Best Local Similarity 63.6%; Pred. No. 17;
XX Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
XX
XX QY 4 MPVLVPEPPFF 14
XX :|||||
XX Db 23 LPVLQSPFFY 33
XX
XX Search completed: April 12, 2005, 08:18:21
XX Job time : 86 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 12, 2005, 07:49:30 ; Search time 17.5909 Seconds
(without alignments)
63.654 Million cell updates/sec

Title: US-10-691-330-5

Perfect score: 81

Sequence: 1 DLEMPVLVPEPPFV 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA.*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	81	100.0	15	4	US-09-641-803-5
2	45	55.6	803	4	US-09-252-991A-30479
3	44	54.3	259	4	US-09-328-352-4873
4	42	51.9	148	4	US-09-270-767-56845
5	42	51.9	218	4	US-09-538-092-178
6	42	51.9	242	4	US-09-252-991A-28361
7	42	51.9	261	4	US-09-270-767-59826
8	42	51.9	330	4	US-09-270-767-41611
9	42	51.9	453	4	US-09-270-767-44397
10	42	51.9	467	4	US-09-270-767-61831
11	42	51.9	704	4	US-09-270-767-46262
12	41.5	51.2	78	4	US-09-513-999C-5875
13	41.5	51.2	86	4	US-09-461-325-456
14	41.5	51.2	86	4	US-10-012-542-456
15	41.5	51.2	86	4	US-10-115-123-456
16	41	50.6	264	4	US-09-902-540-15704
17	41	50.6	1040	4	US-09-902-540-11105
18	41	50.6	1270	4	US-09-248-796A-15522
19	40	49.4	143	4	US-09-621-976-5226
20	40	49.4	219	3	US-09-527-345-2
21	40	49.4	219	4	US-09-923-236-2
22	40	49.4	577	4	US-09-489-039A-9575
23	39.5	48.8	459	4	US-09-902-540-10492
24	39.5	48.8	651	4	US-09-932-678-2
25	39	48.1	220	4	US-09-198-452A-211
26	39	48.1	220	4	US-09-438-185A-194
27	39	48.1	263	4	US-09-902-540-15616

Sequence 374, Appl
Sequence 22, Appl
Sequence 22, Appl
Sequence 1299, Ap
Sequence 20775, A
Sequence 6, Appli
Sequence 8, Appli
Sequence 15517, A
Sequence 2, Appli
Sequence 16293, A
Sequence 5498, Ap
Sequence 14914, A
Sequence 40882, A
Sequence 56098, A
Sequence 19489, A
Sequence 17298, A
Sequence 7220, Ap
Sequence 7221, Ap

ALIGNMENTS

RESULT 1

US-09-641-803-5

; Sequence 5, Application US/09641803

; Patent No. 6500798

; GENERAL INFORMATION:

; APPLICANT: STANTON, G. John

; APPLICANT: HUGHES, Thomas K.

; APPLICANT: BOLDOGH, Istvan

; TITLE OF INVENTION: USE OF COLOSTRININ, CONSTITUENT PEPTIDES THEREOF, AND ANALOGS THEREOF AS OXIDATIVE STRESS REGULATORS

; FILE REFERENCE: 265.00220101

; CURRENT APPLICATION NUMBER: US/09/641.803

; CURRENT FILING DATE: 2000-08-17

; PRIOR APPLICATION NUMBER: 60/149,310

; PRIOR FILING DATE: 1999-08-17

; NUMBER OF SEQ ID NOS: 34

; SOFTWARE: Patentin Ver. 2.1

; SEQ ID NO 5

; LENGTH: 15

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: synthetic

; OTHER INFORMATION: peptide

US-09-641-803-5

Query Match 100.0%; Score 81; DB 4; Length 15;

Best Local Similarity 100.0%; Pred. No. 3.9e-06;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DLEMPVLVPEPPFV 15

|||||

Db 1 DLEMPVLVPEPPFV 15

RESULT 2

US-09-252-991A-30479

; Sequence 30479, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252.991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 30479
; LENGTH: 803
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30479

Query Match 55.6%; Score 45; DB 4; Length 803;
Best Local Similarity 53.3%; Pred. No. 69;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 DLEMPVLVPEPPFV 15
| | | | | | | | | | | | | | | | | | | | |
Db 318 DAEPFVPEVQLPV 332

RESULT 3
US-09-328-352-4873
; Sequence 4873, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 4873
; LENGTH: 259
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-4873

Query Match 54.3%; Score 44; DB 4; Length 259;
Best Local Similarity 58.3%; Pred. No. 30;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 4 MPVLVPEPPFV 15
| | | | | | | | | | | | | | | | | | | | |
Db 166 MVVRPVPYPI 177

RESULT 4
US-09-270-767-56845
; Sequence 56845, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of *Drosophila melanogaster*
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 56845
; LENGTH: 148
; TYPE: PRT
; ORGANISM: *Drosophila melanogaster*
US-09-270-767-56845

Query Match 51.9%; Score 42; DB 4; Length 148;
Best Local Similarity 42.9%; Pred. No. 33;
Matches 6; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 DLEMPVLVPEPPF 14
| | | | | | | | | | | | | | | | | | | | |
Db 6 DAQCPAVFMAWPF 19

RESULT 5
US-09-538-092-178
; Sequence 178, Application US/09538092

; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Giot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CuratPatSeqFormat Version 0.9
; SEQ ID NO 178
; LENGTH: 218
; TYPE: PRT
; ORGANISM: *Saccharomyces cerevisiae*
; FEATURES:
; NAME/KEY: misc feature
; LOCATION: (0)..(0)
; OTHER INFORMATION: Polypeptide Accession Number YDR236C
US-09-538-092-178

Query Match 51.9%; Score 42; DB 4; Length 218;
Best Local Similarity 53.3%; Pred. No. 50;
Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 DLEMPVLVPEPPFV 15
| | | | | | | | | | | | | | | | | | | | |
Db 37 DLPIPAQGPFPV 51

RESULT 6
US-09-252-991A-28361
; Sequence 28361, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 28361
; LENGTH: 242
; TYPE: PRT
; ORGANISM: *Pseudomonas aeruginosa*
US-09-252-991A-28361

Query Match 51.9%; Score 42; DB 4; Length 242;
Best Local Similarity 58.3%; Pred. No. 55;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 3 EMPVLVPEPPF 14
| | | | | | | | | | | | | | | | | | | | |
Db 60 ESPRPAQPPF 71

RESULT 7
US-09-270-767-59826
; Sequence 59826, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of *Drosophila melanogaster*
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767

```
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 59826
; LENGTH: 261
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-59826

Query Match          51.9%; Score 42; DB 4; Length 261;
Best Local Similarity 42.9%; Pred. No. 60;
Matches 6; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 DLEMPVLVEPPFP 14
   | : | : | : |
Db 119 DAQQPAVPMMPF 132

RESULT 8
US-09-270-767-41611
; Sequence 41611, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 41611
; LENGTH: 330
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-41611

Query Match          51.9%; Score 42; DB 4; Length 330;
Best Local Similarity 42.9%; Pred. No. 77;
Matches 6; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 DLEMPVLVEPPFP 14
   | : | : | : |
Db 188 DAQQPAVPMMPF 201

RESULT 9
US-09-270-767-44397
; Sequence 44397, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 44397
; LENGTH: 453
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-44397

Query Match          51.9%; Score 42; DB 4; Length 453;
Best Local Similarity 42.9%; Pred. No. 1.1e+02;
Matches 6; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 DLEMPVLVEPPFP 14
   | : | : | : |
Db 311 DAQQPAVPMMPF 324

RESULT 10
US-09-270-767-41831
; Sequence 41831, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 41831
; LENGTH: 467
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-61831

Query Match          51.9%; Score 42; DB 4; Length 467;
Best Local Similarity 42.9%; Pred. No. 1.1e+02;
Matches 6; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 DLEMPVLVEPPFP 14
   | : | : | : |
Db 135 DAQQPAVPMMPF 148

RESULT 11
US-09-270-767-46262
; Sequence 46262, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 46262
; LENGTH: 704
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-46262

Query Match          51.9%; Score 42; DB 4; Length 704;
Best Local Similarity 42.9%; Pred. No. 1.7e+02;
Matches 6; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 DLEMPVLVEPPFP 14
   | : | : | : |
Db 372 DAQQPAVPMMPF 385

RESULT 12
US-09-513-999C-5875
; Sequence 5875, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59, US2, REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
```

```
; SOFTWARE: Patent.pm
; SEQ ID NO 5875
; LENGTH: 78
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-513-999C-5875

Query Match      51.2%; Score 41.5; DB 4; Length 78;
Best Local Similarity 47.1%; Pred. No. 20;
Matches 8; Conservative 3; Mismatches 1; Indels 5; Gaps 1;

Qy      2 LEMPVLP-----VEPPFP 13
Db      17 LEVPILPTHLLIHPPF 33

RESULT 13
US-09-461-325-456
; Sequence 456, Application US/09461325A
; Patent No. 6475753
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 94 Human Secreted Proteins
; FILE REFERENCE: P2029P1
; CURRENT APPLICATION NUMBER: US/09/461,325A
; CURRENT FILING DATE: 1999-12-14
; EARLIER APPLICATION NUMBER: PCT/US99/13418
; EARLIER FILING DATE: 1999-06-15
; EARLIER APPLICATION NUMBER: 60/089,507
; EARLIER FILING DATE: 1998-06-16
; EARLIER APPLICATION NUMBER: 60/089,508
; EARLIER FILING DATE: 1998-06-16
; EARLIER APPLICATION NUMBER: 60/089,509
; EARLIER FILING DATE: 1998-06-16
; EARLIER APPLICATION NUMBER: 60/089,510
; EARLIER FILING DATE: 1998-06-16
; EARLIER APPLICATION NUMBER: 60/090,112
; EARLIER FILING DATE: 1998-06-22
; EARLIER APPLICATION NUMBER: 60/090,113
; EARLIER FILING DATE: 1998-06-22
; NUMBER OF SEQ ID NOS: 532
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 456
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-461-325-456

Query Match      51.2%; Score 41.5; DB 4; Length 86;
Best Local Similarity 47.1%; Pred. No. 22;
Matches 8; Conservative 3; Mismatches 1; Indels 5; Gaps 1;

Qy      2 LEMPVLP-----VEPPFP 13
Db      17 LEVPILPTHLLIHPPF 33

RESULT 14
US-10-012-542-456
; Sequence 456, Application US/10012542
; Patent No. 6627741
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 94 Human Secreted Proteins
; FILE REFERENCE: P2029P1
; CURRENT APPLICATION NUMBER: US/10/012,542
; CURRENT FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/461,325
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-12-14
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,507
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,508
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16

; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 456
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-012-542-456

Query Match      51.2%; Score 41.5; DB 4; Length 86;
Best Local Similarity 47.1%; Pred. No. 22;
Matches 8; Conservative 3; Mismatches 1; Indels 5; Gaps 1;

Qy      2 LEMPVLP-----VEPPFP 13
Db      17 LEVPILPTHLLIHPPF 33

RESULT 15
US-10-115-123-456
; Sequence 456, Application US/10115123
; Patent No. 6774216
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 94 Human Secreted Proteins
; FILE REFERENCE: P2029G30APID2
; CURRENT APPLICATION NUMBER: US/10/115,123
; CURRENT FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: PCT/US99/13418
; PRIOR FILING DATE: 1999-06-15
; PRIOR APPLICATION NUMBER: 60/089,507
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089,508
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089,509
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089,510
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/090,112
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090,113
; PRIOR FILING DATE: 1998-06-22
; NUMBER OF SEQ ID NOS: 532
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 456
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-115-123-456

Query Match      51.2%; Score 41.5; DB 4; Length 86;
Best Local Similarity 47.1%; Pred. No. 22;
Matches 8; Conservative 3; Mismatches 1; Indels 5; Gaps 1;

Qy      2 LEMPVLP-----VEPPFP 13
Db      17 LEVPILPTHLLIHPPF 33

Search completed: April 12, 2005, 08:07:52
Job time : 18.5909 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 12, 2005, 07:49:30 ; Search time 46.5 Seconds
(without alignments)
107.096 Million cell updates/sec

Title: US-10-691-330-5

Perfect score: 81

Sequence: 1 DLEMPVLVPEPPFV 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1418010 seqs, 331937259 residues

Total number of hits satisfying chosen parameters: 1418010

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Published Applications AA:*
1: /cgn2_6/ptodata/1/pubaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/1/pubaa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/1/pubaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	81	100.0	15	14 US-10-281-652-5	Sequence 5, Appli
2	81	100.0	15	17 US-10-691-330-5	Sequence 5, Appli
3	57	70.4	10	16 US-10-182-110-3	Sequence 3, Appli
4	49	60.5	104	15 US-10-425-114-58864	Sequence 58864, A
5	48	59.3	232	15 US-10-104-047-3111	Sequence 3111, Ap
6	45	55.6	446	15 US-10-282-122A-77182	Sequence 77182, A
7	44	54.3	135	16 US-10-437-963-108040	Sequence 108040,
8	43	53.1	325	15 US-10-424-599-254766	Sequence 254766,
9	42	51.9	68	15 US-10-424-599-284024	Sequence 284024,
10	42	51.9	174	15 US-10-424-599-204286	Sequence 204286,
11	42	51.9	249	15 US-10-424-599-144697	Sequence 144697,
12	42	51.9	457	15 US-10-369-493-18781	Sequence 18781, A
13	41.5	51.2	86	14 US-10-012-542-4556	Sequence 456, App

14	41.5	51.2	86	14	US-10-115-123-456	Sequence 456, App
15	41	50.6	88	15	US-10-424-599-204279	Sequence 204279,
16	41	50.6	101	15	US-10-276-774-1842	Sequence 1842, Ap
17	41	50.6	113	15	US-10-424-599-223803	Sequence 223803,
18	41	50.6	173	15	US-10-424-599-204281	Sequence 204281,
19	41	50.6	175	15	US-10-424-599-204284	Sequence 204284,
20	41	50.6	175	15	US-10-424-599-204285	Sequence 204285,
21	41	50.6	189	16	US-10-437-963-144895	Sequence 144895,
22	41	50.6	308	16	US-10-437-963-171233	Sequence 171233,
23	41	50.6	397	14	US-10-156-761-8560	Sequence 8560, Ap
24	41	50.6	632	16	US-10-437-963-145459	Sequence 145459,
25	41	50.6	644	16	US-10-437-963-112405	Sequence 112405,
26	41	50.6	8026	14	US-10-132-134-12	Sequence 12, Appli
27	40	49.4	97	16	US-10-767-701-32203	Sequence 32203, A
28	40	49.4	97	15	US-10-424-599-216454	Sequence 216454, A
29	40	49.4	151	15	US-10-425-114-38710	Sequence 38710, A
30	40	49.4	151	16	US-10-437-963-166092	Sequence 166092,
31	40	49.4	167	16	US-10-437-963-192065	Sequence 192065,
32	40	49.4	219	9	US-09-922-480-2	Sequence 2, Appli
33	40	49.4	219	9	US-09-923-236-2	Sequence 2, Appli
34	40	49.4	219	9	US-09-922-469-2	Sequence 2, Appli
35	40	49.4	219	10	US-09-746-783-106	Sequence 106, App
36	40	49.4	219	17	US-10-874-706-29	Sequence 29, Appl
37	40	49.4	219	17	US-10-969-164-2	Sequence 2, Appli
38	40	49.4	221	14	US-10-231-417-192	Sequence 192, App
39	40	49.4	239	15	US-10-424-599-216458	Sequence 216458,
40	40	49.4	264	16	US-10-437-963-104274	Sequence 104274,
41	40	49.4	318	16	US-10-437-963-137718	Sequence 137718,
42	40	49.4	331	15	US-10-425-114-42069	Sequence 42069, A
43	40	49.4	460	15	US-10-425-114-52303	Sequence 52303, A
44	40	49.4	561	15	US-10-424-599-239201	Sequence 239201,
45	40	49.4	850	15	US-10-369-493-10911	Sequence 10911, A

ALIGNMENTS

RESULT 1

US-10-281-652-5
; Sequence 5, Application US/10281652
; Publication No. US20030091606A1
; GENERAL INFORMATION:
; APPLICANT: STANTON, G. John
; APPLICANT: HUGHES, Thomas K.
; APPLICANT: BOLDOGH, Istvan
; TITLE OF INVENTION: USE OF COLOSTRININ, CONSTITUENT PEPTIDES THEREOF, AND ANALOGS THEREOF AS OXIDATIVE STRESS REGULATORS
; FILE REFERENCE: 265.00220101
; CURRENT APPLICATION NUMBER: US/10/281.652
; CURRENT FILING DATE: 2002-10-28
; PRIOR FILING DATE: 2000-08-17
; PRIOR APPLICATION NUMBER: 60/149,310
; PRIOR FILING DATE: 1999-08-17
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic peptide
; OTHER INFORMATION: peptide
US-10-281-652-5

Query Match 100.0%; Score 81; DB 14; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.9e-05;
Matches: 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 DLEMPVLVPEPPFV 15
Db 1 DLEMPVLVPEPPFV 15

RESULT 2
US-10-691-330-5
; Sequence 5, Application US/10691330
; Publication No. US20050042300A1
; GENERAL INFORMATION:
; APPLICANT: Boldogh, Istvan
; APPLICANT: Stanton, G. John
; APPLICANT: Georgiades, Jerry A.
; APPLICANT: Hughes, Thomas K., Jr.
; APPLICANT: Kruzel, Marian
; TITLE OF INVENTION: USE OF COLOSTRININ, CONSTITUENT PEPTIDES THEREOF, AND ANALOGS
; FILE REFERENCE: 265.00390101
; CURRENT APPLICATION NUMBER: US/10/691,330
; CURRENT FILING DATE: 2003-10-22
; PRIOR APPLICATION NUMBER: US 60/420,369
; PRIOR FILING DATE: 2002-10-22
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 5
; LENGTH: 15
; TYPE: PRT
; ORGANISM: artificial
; FEATURE:
; OTHER INFORMATION: Synthetic Peptides
US-10-691-330-5

Query Match 100.0%; Score 81; DB 17; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.9e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DLEMPVLVEPPFV 15
Db 1 DLEMPVLVEPPFV 15

RESULT 3
US-10-182-110-3
; Sequence 3, Application US/10182110
; Publication No. US20040171553A1
; GENERAL INFORMATION:
; APPLICANT: Regen Therapeutics plc
; TITLE OF INVENTION: Peptide Fragments of Colostrinin And Their Use
; FILE REFERENCE: AAT-14173
; CURRENT APPLICATION NUMBER: US/10/182,110
; CURRENT FILING DATE: 2003-04-21
; PRIOR APPLICATION NUMBER: GB0001825.9
; PRIOR FILING DATE: 2000-01-26
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 3
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Truncated version of a peptide found in colostrinin
US-10-182-110-3

Query Match 70.4%; Score 57; DB 16; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.047;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 PVLVEPPFV 14
Db 1 PVLVEPPFV 10

RESULT 4
US-10-425-114-58864
; Sequence 58864, Application US/10425114
; Publication No. US20040034888A1

; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 58864
; LENGTH: 104
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700431848_FLI.pep
US-10-425-114-58864

Query Match 60.5%; Score 49; DB 15; Length 104;
Best Local Similarity 60.0%; Pred. No. 8;
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 DLEMPVLVEPPFV 15
Db 61 DLEMPVLVEPPFV 75

RESULT 5
US-10-104-047-3111
; Sequence 3111, Application US/10104047
; Publication No. US20030236392A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20030236392A1 full length cDNA
; FILE REFERENCE: HI-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 3111
; LENGTH: 232
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-104-047-3111

Query Match 59.3%; Score 48; DB 15; Length 232;
Best Local Similarity 66.7%; Pred. No. 26;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 3 EMPVLVEPPFV 14
Db 209 KFPVLVEPPFV 220

RESULT 6
US-10-282-122A-77182
; Sequence 77182, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant

APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: Patent in version 3.1
SEQ ID NO 77182
LENGTH: 446
TYPE: PRT
ORGANISM: Vibrio cholerae
US-10-282-122A-77182

Query Match 55.6%; Score 45; DB 15; Length 446;
Best Local Similarity 53.3%; Pred. No. 1.4e+02;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 1 DLEMPVLVPEPPFV 15
|||:|:|:|:
Db 76 DLEVPNLEAEPYL 90

RESULT 7
US-10-437-963-108040
Sequence 108040, Application US/10437963
Publication No. US20040123343A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 108040
LENGTH: 135
TYPE: PRT
ORGANISM: Oryza sativa
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_12332C.1.ppep
US-10-437-963-108040

Query Match 54.3%; Score 44; DB 16; Length 135;

Best Local Similarity 47.4%; Pred. No. 58;
Matches 9; Conservative 2; Mismatches 4; Indels 4; Gaps 1;
Qy 1 DLEMPVLVPEP---PPFV 15
|||:|:|:|:
Db 75 DAPMPEIPIHPPPVPPFV 93
RESULT 8
US-10-424-599-254766
Sequence 254766, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 254766
LENGTH: 325
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_72076C.1.ppep
US-10-424-599-254766

Query Match 53.1%; Score 43; DB 15; Length 325;
Best Local Similarity 77.8%; Pred. No. 2e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 5 PVLPEPFP 13
|||:|:|:|:
Db 191 PVPVHPFP 199

RESULT 9
US-10-424-599-284024
Sequence 284024, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 284024
LENGTH: 68
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_9849C.1.ppep
US-10-424-599-284024

Query Match 51.9%; Score 42; DB 15; Length 68;
Best Local Similarity 53.3%; Pred. No. 57;
Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 1 DLEMPVLVPEPPFV 15
|||:|:|:|:
Db 44 DAEPCLYKEPPLI 58

RESULT 10

US-10-424-599-204286
; Sequence 204286, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 204286
; LENGTH: 174
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_26498C.1.pep
US-10-424-599-204286

Query Match 51.9%; Score 42; DB 15; Length 174;
Best Local Similarity 54.5%; Pred. No. 1.5e+02;
Matches 6; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DLEMPVLVPEP 11
:::|||||
Db 77 NIDIPVLPVEP 87

RESULT 11

US-10-424-599-144697
; Sequence 144697, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 144697
; LENGTH: 249
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_101680C.1.pep
US-10-424-599-144697

Query Match 51.9%; Score 42; DB 15; Length 249;
Best Local Similarity 80.0%; Pred. No. 2.2e+02;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 5 PVLVPVEPPFP 14
|||||
Db 48 PVLVPVEGLPF 57

RESULT 12

US-10-369-493-18781
; Sequence 18781, Application US/10369493
; Publication No. US2003023675A1
; GENERAL INFORMATION:
; APPLICANT: Cao Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng

; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 18781
; LENGTH: 457
; TYPE: PRT
; ORGANISM: Anabaena PCC7120
US-10-369-493-18781

Query Match 51.9%; Score 42; DB 15; Length 457;
Best Local Similarity 63.6%; Pred. No. 4e+02;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 5 PVLVPVEPPFPV 15
:::|||||
Db 364 PIMPVMPFPDV 374

RESULT 13

US-10-012-542-456
; Sequence 456, Application US/10012542
; Publication No. US20030044851A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 94 Human Secreted Proteins
; FILE REFERENCE: PZ029P1
; CURRENT APPLICATION NUMBER: US/10/012,542
; CURRENT FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/461,325
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-12-14
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,507
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,508
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,509
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,510
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/090,112
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/090,113
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-22
; NUMBER OF SEQ ID NOS: 532
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 456
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-012-542-456

Query Match 51.2%; Score 41.5; DB 14; Length 86;
Best Local Similarity 47.1%; Pred. No. 86;
Matches 8; Conservative 3; Mismatches 1; Indels 5; Gaps 1;

Qy 2 LEMPVLP-----VEPPFP 13
||:|:|:|:|
Db 17 LEVPILPHTLLTHPFP 33

RESULT 14

US-10-115-123-456
; Sequence 456, Application US/10115123
; Publication No. US20030065151A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 94 Human Secreted Proteins
; FILE REFERENCE: PZ029G30AP1D2
; CURRENT APPLICATION NUMBER: US/10/115,123

```

; CURRENT FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: PCT/US99/13418
; FILING DATE: 1999-06-15
; PRIOR APPLICATION NUMBER: 60/089,507
; FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089,508
; FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089,509
; FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089,510
; FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/090,112
; FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090,113
; FILING DATE: 1998-06-22
; NUMBER OF SEQ ID NOS: 532
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 456
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-115-123-456

Query Match      51.2%; Score 41.5; DB 14; Length 86;
Best Local Similarity 47.1%; Pred. No. 86;
Matches 8; Conservative 3; Mismatches 1; Indels 5; Gaps 1;

QY 2 LEMPVLP-----VERFP 13
   ||:|:|
DB 17 LEVPILPTHLLIHFP 33

RESULT 15
US-10-424-599-204279
; Sequence 204279, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 204279
; LENGTH: 88
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(88)
; OTHER INFORMATION: unsure at all xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_26491C.1.pep
US-10-424-599-204279

Query Match      50.6%; Score 41; DB 15; Length 88;
Best Local Similarity 45.5%; Pred. No. 1e+02;
Matches 5; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLEMPVLPVEP 11
   :|:|:|:|
DB 77 NIDIEVPIEP 87

```

Search completed: April 12, 2005, 07:55:21
Job time : 46.5 secs

This Page Blank (uspto)

Best Local Similarity 53.3%; Pred. No. 15;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 DLEMPVLVPVPPFV 15
Db 76 DLEVPNLELAFPL 90

RESULT 3
H69490
formylmethanofuran dehydrogenase (tungsten) chain B homolog (fwdB-2) - Archaeoglobus fulgidus
C:Species: Archaeoglobus fulgidus
C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C:Accession: H69490
R;Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
R;Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A:Authors: Uterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.
Smith, H.O.; Woese, C.R.; Venter, J.C.
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeon
A:Reference number: A69250; MUID:98049343; PMID:9389475
A:Accession: H69490
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-421 <KLE>
A:Cross-references: UNIPROT:Q28350; GB:AE000970; GB:AE000782; NID:Q2669293; PIDN:AB8932
C:Superfamily: formylmethanofuran dehydrogenase (molybdenum) beta chain
C:Keywords: iron-sulfur protein; metalloprotein; molybdenum; molybdopterin

Query Match 53.1%; Score 43; DB 2; Length 421;
Best Local Similarity 54.5%; Pred. No. 29;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 3 EMPVLVPVPPFV 13
Db 350 EIPVIQIDPPF 360

RESULT 4
T24690
hypothetical protein T08D2.8 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T24690
R;McMurray, A.
submitted to the EMBL Data Library, March 1997
A:Reference number: Z19924
A:Accession: T24690
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-556 <WIL>
A:Cross-references: UNIPROT:Q9XTX7; EMBL:Z92839; PIDN:CAB07415.1; CESP:T08D2.8
A:Experimental source: clone T08D2
C:Genetics:
A:Gene: CESP:T08D2.8
A:Introns: 5/1; 54/3; 80/3; 255/3; 307/3; 352/3; 479/1

Query Match 53.1%; Score 43; DB 2; Length 556;
Best Local Similarity 72.7%; Pred. No. 40;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 5 FVLVPVPPFV 15
Db 443 PVAPVKPFPFV 453

RESULT 5
S54532
probable membrane protein YDR236c - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein YD8419.03c
C:Species: Saccharomyces cerevisiae
C>Date: 08-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 09-Jul-2004

C:Accession: S54532
R;Oliver, K.; Harris, D.
submitted to the EMBL Data Library, May 1995
A:Reference number: S54530
A:Accession: S54532
A:Molecule type: DNA
A:Residues: 1-218 <OLI>
A:Cross-references: UNIPROT:Q03778; EMBL:Z49701; NID:g817819; PID:g817822; GSPDB:GN000004
A:Experimental source: strain AB972
C:Genetics:
A:Gene: SGD:FMN1; MIPS:YDR236c
A:Cross-references: SGD:S0002644
A:Map position: 4R
C:Keywords: transmembrane protein
F;4-20/Domain: transmembrane #status predicted <TMM>

Query Match 51.9%; Score 42; DB 2; Length 218;
Best Local Similarity 53.3%; Pred. No. 20;
Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 DLEMPVLVPVPPFV 15
Db 37 DLPIPAQPGPPPLV 51

RESULT 6
JC6530
laminin receptor processed pseudogene LAMRL5 - human
C:Species: Homo sapiens (man)
C>Date: 28-Aug-1998 #sequence_revision 28-Aug-1998 #text_change 28-Aug-1998
C:Accession: JC6530
R;Richardson, M.P.; Braybrook, C.; Tham, M.; Moore, G.E.; Stanier, P.
Gene 266, 145-150, 1998
A:Title: Molecular cloning and characterization of a highly conserved human 67-kDa laminin
A:Reference number: JC6530; MUID:98121324; PMID:9461426
A:Accession: JC6530
A>Status: conceptual translation of pseudogene
A:Molecule type: DNA
A:Residues: 1-295 <RIC>
A:Experimental source: brain
C:Comment: No evidence could be found that this intronless gene sequence is expressed.
C:Genetics:
A:Gene: LAMRL5
A:Map position: Xg21.3
A:Introns: #status absent
C:Keywords: brain; glycoprotein; laminin binding; pseudogene; receptor

Query Match 51.9%; Score 42; DB 4; Length 295;
Best Local Similarity 46.2%; Pred. No. 28;
Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 DLEMPVLVPVPPFV 13
Db 251 EVEVPSPFIEFPF 263

RESULT 7
T24298
hypothetical protein T01E8.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T24298
R;McMurray, A.
submitted to the EMBL Data Library, March 1995
A:Reference number: Z19871
A:Accession: T24298
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-367 <WIL>
A:Cross-references: UNIPROT:Q22069; EMBL:Z48809; PIDN:CAA88744.1; GSPDB:GN000020; CESP:T01E8.2
A:Experimental source: clone T01E8
C:Genetics:
A:Gene: CESP:T01E8.2

A;Map position: 2
A;Introns: 48/2; 200/2; 254/3

Query Match 51.9%; Score 42; DB 2; Length 367;
Best Local Similarity 60.0%; Pred. No. 36;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 5 PVLVPVEPPFV 14
:|:|:|:|:|
Db 295 PIFPFRPPFV 304

RESULT 8

AF1942
hypothetical protein all1089 [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C;Accession: AF1942
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AF1942
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-448 <KUR>
A;Cross-references: UNIPROT:Q8YXW8; GB:BA000019; PIDN:BA073046.1; PID:g17130435; GSPDB:G
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: all1089

Query Match 51.9%; Score 42; DB 2; Length 448;
Best Local Similarity 60.0%; Pred. No. 46;
Matches 6; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 3 EMPVLVPEPFV 12
|:|:|:|:|:|
Db 98 ELPLLPVDPI 107

RESULT 9

AF2433
aldehyde dehydrogenase [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C;Accession: AF2433
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AF2433
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-473 <KUR>
A;Cross-references: UNIPROT:Q8YMB2; GB:BA000019; PIDN:BA076721.1; PID:g17134160; GSPDB:G
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: all1022
C;Superfamily: NAD-dependent aldehyde dehydrogenase; aldehyde dehydrogenase homology

Query Match 51.9%; Score 42; DB 2; Length 473;
Best Local Similarity 63.6%; Pred. No. 49;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 5 PVLVPVEPPFV 15
|:|:|:|:|:|
Db 365 PIMPVMPDPV 375

RESULT 10

T07447
DNA-directed RNA polymerase (EC 2.7.7.6) beta'-1 chain - Japanese black pine chloroplast
C;Species: chloroplast Pinus thunbergiana (Japanese black pine)
C;Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 09-Jul-2004
C;Accession: T07447
R;Wakaugi, T.; Tsudzuki, J.; Ito, S.; Nakashima, K.; Tsudzuki, T.; Sugiyama, M.
Proc. Natl. Acad. Sci. U.S.A. 91, 9794-9798, 1994
A;Title: Loss of all ndh genes as determined by sequencing the entire chloroplast genome
A;Reference number: Z16030; MUID:95024047; PMID:7937893
A;Accession: T07447
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-696 <WAK>
A;Cross-references: UNIPROT:P52733; EMBL:D17510; NID:g529643; PIDN:BA023472.1; PID:g52626;
C;Genetics:

A;Gene: rpoC1
A;Genome: chloroplast
A;Note: intron positions not resolved (incomplete sequence)
C;Superfamily: chloroplast DNA-directed RNA polymerase beta'-1 chain
C;Keywords: chloroplast; nucleotidyltransferase; transcription

Query Match 51.9%; Score 42; DB 2; Length 696;
Best Local Similarity 66.7%; Pred. No. 75;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 4 MPVLVPEPPFV 15
:|:|:|:|:|
Db 285 LPVLPPPEPPV 296

RESULT 11

T17426
FK506 polyketide synthetase fkbp [imported] - Streptomyces sp. (strain MA6548)
C;Species: Streptomyces sp.
A;Variety: strain MA6548
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T17426
R;Motamedi, H.; Shafiee, A.
Eur. J. Biochem. 256, 528-534, 1998
A;Title: The biosynthetic gene cluster for the macrolactone ring of the immunosuppressant FK506
A;Reference number: Z18779; MUID:98451508; PMID:9780228
A;Accession: T17426
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-1504 <NOT>
A;Cross-references: UNIPROT:Q9ZGAG; EMBL:AF082100; NID:g3798623; PID:g3798625; PIDN:AAC6
A;Experimental source: strain MA6548
C;Genetics:

A;Gene: fkbp
C;Function:
A;Description: required during the biosynthesis of the immunosuppressant FK506 for the
C;Superfamily: Mycobacterium tuberculosis mbeE protein; acetate-CoA ligase homology; acy
C;Keywords: carrier protein; phosphopantetheine; phosphoprotein
F;533-982/Domain: acetate-CoA ligase homology <ACL>
F;999-1067/Domain: acyl carrier protein homology <ACP>
F;1031/Binding site: phosphopantetheine (Ser) (covalent) #status predicted

Query Match 51.9%; Score 42; DB 2; Length 1504;
Best Local Similarity 53.3%; Pred. No. 1.8e+02;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 1 DLEMPVLVPEPPFV 15
:|:|:|:|:|
Db 1423 ELRLPGLRTEPPFV 1437

RESULT 12

B97432
glucose-6-phosphate 1-dehydrogenase (96pd) [imported] - Agrobacterium tumefaciens (strain
C;Species: Agrobacterium tumefaciens
C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
C;Accession: B97432

R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldman, A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.; Science 294, 2323-2328, 2001

A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent *Agrobacterium tumefaciens* 294, 2323-2328, 2001

A;Reference number: A97359; MUID:21608551; PMID:11743194

A;Accession: B97432

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-503 <KUR>

A;Cross-references: UNIPROT:Q8UHS9; GB:AE007869; PIDN:AAK86411.1; PID:gl5155545; GSPDB:G

C;Genetics:

A;Gene: AGR_C.1065

A;Map position: circular chromosome

C;Superfamily: glucose-6-phosphate dehydrogenase

Query Match 50.6%; Score 41; DB 2; Length 503;

Best Local Similarity 46.7%; Pred. No. 75;

Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 DLEMPVLVPEPPFV 15

Db 12 DMSSQIIPVEFFDCV 26

RESULT 13

T41201

isoleucyl-trna synthetase - fission yeast (*Schizosaccharomyces pombe*)

C;Species: *Schizosaccharomyces pombe*

C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004

C;Accession: T41201

R;Wood, V.; Rajandream, M.A.; Barrell, B.G.; Jimenez Martinez, J.

submitted to the EMBL Data Library, July 1999

A;Reference number: Z21978

A;Accession: T41201

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Residues: 1-973 <WOO>

A;Cross-references: UNIPROT:Q9HSL3; EMBL:AL109726; PIDN:CAN52155.1; GSPDB:GN00068

C;Genetics:

A;Gene: SPC18B5.08c

A;Map position: 3

C;Superfamily: isoleucine-trna ligase

Query Match 50.6%; Score 41; DB 2; Length 973;

Best Local Similarity 53.8%; Pred. No. 1.6e+02;

Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 3 EMPVLVPEPPFV 15

Db 325 ENLLPKQSPFFL 337

RESULT 14

T00333

hypothetical protein KIAA0560 - human

C;Species: *Homo sapiens* (man)

C;Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 09-Jul-2004

C;Accession: T00333

R;Nagase, T.; Ishikawa, K.; Miyajima, N.; Tanaka, A.; Kotani, H.; Nomura, N.; Ohara, O.

DNA Res. 5, 31-39, 1998

A;Title: Prediction of the coding sequences of unidentified human genes. IX. The complete

A;Reference number: Z14086; MUID:98290545; PMID:9628581

A;Accession: T00333

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-1421 <NAG>

A;Cross-references: UNIPROT:O60306; EMBL:AB011132; NID:d1185402; PIDN:BAA25486.1

A;Experimental source: brain; clone H14648

C;Genetics:

A;Note: KIAA0560

Query Match 50.6%; Score 41; DB 2; Length 1421;

Best Local Similarity 50.0%; Pred. No. 2.4e+02;

Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 2 LEMPVLVPEPPF 13

Db 1264 LHLHIPTBPPF 1275

RESULT 15

AG3000

1-acyl-sn-glycerol-3-phosphate acyltransferase p1aC [imported] - *Agrobacterium tumefaciens*

C;Species: *Agrobacterium tumefaciens*

C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 16-Aug-2004

C;Accession: AG3000

R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.

erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClellan,

Karp, P.; Romero, P.; Zhang, S.

Science 294, 2317-2323, 2001

A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, E.

ster, E.W.

A;Title: The Genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58.

A;Reference number: AB2577; MUID:21608550; PMID:11743193

A;Accession: AG3000

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-266 <KUR>

A;Cross-references: UNIPROT:Q8U9H6; GB:AE008689; PIDN:AAL44421.1; PID:gl7742021; GSPDB:G

A;Experimental source: strain C58 (Dupont)

C;Genetics:

A;Gene: p1aC

A;Map position: linear chromosome

C;Superfamily: 1-acyl-sn-glycerol-3-phosphate acyltransferase

Query Match 49.4%; Score 40; DB 2; Length 266;

Best Local Similarity 53.8%; Pred. No. 53;

Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 DLEMPVLVPEPPF 13

Db 174 DLQVPVPMHP 186

Search completed: April 12, 2005, 07:56:53

Job time : 12.7727 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 12, 2005, 07:49:31 ; Search time 71.1818 Seconds
(without alignments)
107.909 Million cell updates/sec

Title: US-10-691-330-5

Perfect score: 81

Sequence: 1 DLEMPVLVPFPFV 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	48	59.3	232	Q8NAJ2	Q8NAJ2 homo sapien
2	47	58.0	222	Q75DY2	Q75DY2 ashbya goss
3	47	58.0	1766	P78584	P78584 aspergillus
4	46	56.8	180	Q7RW43	Q7RW43 methanopyru
5	46	56.8	935	Q7RZ84	Q7RZ84 neurospora
6	46	56.8	948	Q6M999	Q6M999 neurospora
7	45	55.6	79	Q99ME4	Q99ME4 rattus norv
8	45	55.6	80	Q8LF16	Q8LF16 arabidopsis
9	45	55.6	91	Q8SR85	Q8SR85 arabidopsis
10	45	55.6	200	Q6GMA4	Q6GMA4 xenopus lae
11	45	55.6	219	Q22140	Q22140 caenorhabdi
12	45	55.6	446	Q9KTH8	Q9KTH8 vibrio chol
13	45	55.6	557	Q99L21	Q99L21 mus musculu
14	45	55.6	599	Q8C2N3	Q8C2N3 mus musculu
15	45	55.6	599	Q8C3K9	Q8C3K9 mus musculu
16	45	55.6	599	Q8C4X8	Q8C4X8 mus musculu
17	45	55.6	607	Q8BSX2	Q8BSX2 mus musculu
18	45	55.6	619	Q9H5C7	Q9H5C7 homo sapien
19	45	55.6	668	Q9H5C7	Q9H5C7 homo sapien
20	45	55.6	772	Q9BLH1	Q9BLH1 bombyx mori
21	45	55.6	902	Q8G762	Q8G762 bifidobacte
22	45	55.6	1307	Q9C093	Q9C093 homo sapien
23	45	55.6	1744	Q9R095	Q9R095 rattus norv
24	44	54.3	191	Q6ACW3	Q6ACW3 leifsonia.x
25	44	54.3	428	Q6K6Z8	Q6K6Z8 oryza sativ
26	44	54.3	593	Q7RN16	Q7RN16 plasmodium
27	44	54.3	638	Q75RZ6	Q75RZ6 ciona intes
28	44	54.3	656	Q2754Q2	Q2754Q2 ashbya goss
29	44	54.3	726	Q6G327	Q6G327 bartonella
30	43	53.1	388	Q9XTX7	Q9XTX7 caenorhabdi
31	43	53.1	421	Q28350	Q28350 archaeoglob

32	43	53.1	444	2	Q9SR06	Q9SR06 arabidopsis
33	43	53.1	1161	1	KCH2_RABIT	KCH2_RABIT o potassium
34	42.5	52.5	681	2	Q9LRV1	Q9LRV1 arabidopsis
35	42	51.9	157	2	Q7QJ10	Q7QJ10 anopheles g
36	42	51.9	200	2	Q67RZ8	Q67RZ8 symbiobacte
37	42	51.9	218	1	RIFK_YEAST	RIFK_YEAST saccharomyc
38	42	51.9	232	2	Q9VVG4	Q9VVG4 drosophila
39	42	51.9	260	2	Q94AV8	Q94AV8 arabidopsis
40	42	51.9	265	2	Q889N4	Q889N4 pseudomonas
41	42	51.9	367	2	Q22069	Q22069 caenorhabdi
42	42	51.9	448	2	Q8YXW8	Q8YXW8 anabaena sp
43	42	51.9	449	2	Q6AGJ9	Q6AGJ9 leifsonia x
44	42	51.9	468	2	Q9NF32	Q9NF32 drosophila
45	42	51.9	469	2	Q9W5D6	Q9W5D6 drosophila

ALIGNMENTS

RESULT 1
Q8NAJ2 PRELIMINARY; PRT; 232 AA.
AC Q8NAJ2; 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein FLJ35269.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TrEMBLrel. 22, Created;
RX PubMed=14702039; DOI=10.1038/ng1285;
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA Yamamoto J., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahari K.,
RA Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shiratori A.,
RA Sudo H., Hoshiro T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y.,
RA Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M., Yamazaki M.,
RA Ninomiya K., Ishibashi T., Yamashita H., Tanikawa K., Fujimori K.,
RA Tanai H., Kimata M., Watanabe M., Hirakawa S., Chiba Y., Ishida S.,
RA Ono Y., Takiguchi S., Watanabe S., Iosida M., Hosoda T., Kusanagi J.,
RA Kanehori K., Takahashi-Fujii A., Hara H., Tanase T., Nomura Y.,
RA Togiya S., Komai F., Hara R., Takeuchi K., Arita M., Imose N.,
RA Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,
RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,
RA Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
RA Hishigaki H., Watanabe K., Sugiyama A., Takemoto M., Kawakami B.,
RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
FT "Complete sequencing and characterization of 21,243 full-length human
cDNAs.";
RL Nat. Genet. 36:40-45(2004).
DR EMBL: AK092588; BAC03921.1; -;
SQ SEQUENCE 232 AA; 25113 MW; 36707CBA84594AC8 CRC64;

Query Match 59.3%; Score 48; DB 2; Length 232;

Best Local Similarity 66.7%; Pred. No. 12;

Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```

QY      3 EMPVLVPVPPF 14
Db      209 KFPVLVPVHPVPPF 220

RESULT 2
Q75DY2 ID Q75DY2 PRELIMINARY; PRT; 222 AA.
AC Q75DY2
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE ABL109WP.
OS ORFNames=ABL109W;
GN Aehya gossypii (Yeast) (Eremothecium gossypii).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Eremothecium.
OX NCBI_TaxID=33169;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 10895;
RA Voegeli S.E., Dietrich F.S., Brachat S., Lerch A., Gaffney T.,
RA Philippsen P.;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE016815; AAS50662.1; -.
DR AGD; ABL109W; -.
DR GO; GO:0008531; F:riboflavin kinase activity; IEA.
DR GO; GO:0009231; P:vitamin B2 biosynthesis; IEA.
DR InterPro; IPR002606; FAD_Synth.
DR Pfam; PF01687; FAD_Synth; 1.
DR ProDom; PD003662; FAD_Synth; 1.
SQ SEQUENCE 222 AA; 25125 MW; 965B95DD257FB439 CRC64;

Query Match 58.0%; Score 47; DB 2; Length 222;
Best Local Similarity 46.7%; Pred. No. 17;
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY      1 DLEMPVLVPVPPFV 15
Db      43 DIPIPASFVQPPFL 57

RESULT 3
P78584 ID P78584 PRELIMINARY; PRT; 1766 AA.
AC P78584
DT 01-MAY-1997 (TREMBlrel. 03, Created)
DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Polyketide synthase PKS2.
GN Name=pks2;
OS Aspergillus parasiticus.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=5067;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL 2999;
RX MEDLINE=98268975; PubMed=9603849;
RA Feng G.H., Leonard T.J.;
RA "Culture conditions control expression of the genes for aflatoxin and
RT sterigmatocystin biosynthesis in Aspergillus parasiticus and A.
RT nidulans";
RL Appl. Environ. Microbiol. 64:2275-2277 (1998).
DR EMBL; U52151; AAC23536.1; -.
DR GO; GO:0048037; F:cofactor binding; IEA.
DR GO; GO:0016740; P:transferase activity; IEA.
DR GO; GO:0006633; P:fatty acid biosynthesis; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR001227; Ac_transferase.
DR InterPro; IPR000794; Ketoacyl_synth.
DR InterPro; IPR006163; Phosphateteth_bind.
DR Pfam; PF00698; Acyl_transf_1; 1.

DR Pfam; PF00109; ketoacyl-synt; 1.
DR Pfam; PF02801; ketoacyl-synt C; 1.
DR Pfam; PF00550; PP-binding; 1.
DR PROSITE; PS00075; ACP_DOMAIN; 1.
DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; 1.
KW Transferase.
SQ SEQUENCE 1766 AA; 192067 MW; E20C4BF26F60671E CRC64;

Query Match 58.0%; Score 47; DB 2; Length 1766;
Best Local Similarity 60.0%; Pred. No. 1.6e+02;
Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY      1 DLEMPVLVPVPPFV 15
Db      1340 DLEMPVLPLATMKV 1354

RESULT 4
Q8TW43 ID Q8TW43 PRELIMINARY; PRT; 180 AA.
AC Q8TW43
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Uncharacterized membrane protein.
GN OrderedLocusNames=MK1194;
OS Methanopyrus kandleri.
OC Archaea; Euryarchaeota; Methanopyri; Methanopyrales; Methanopyraceae;
OC Methanopyrus.
OX NCBI_TaxID=2320;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AV19 / DSM 6324 / JCM 9639;
RX MEDLINE=21927647; PubMed=11930014; DOI=10.1073/pnas.032671499;
RA Slesarev A.I., Mezhevaya K.V., Makarova K.S., Polushin N.N.,
RA Shcherbinina O.V., Shakhova V.V., Belova G.I., Aravind L.,
RA Natale D.A., Rogozin I.B., Tatusov R.L., Wolf Y.I., Stetter K.O.,
RA Malykh A.G., Koonin E.V., Kozhavkin S.A.;
RT "The complete genome of hyperthermophile Methanopyrus kandleri AV19
RT and monophyly of archaeal methanogens";
RL Proc. Natl. Acad. Sci. U.S.A. 99:4644-4649 (2002).
DR EMBL; AE010411; AAM02407.1; -.
KW Complete proteome.
SQ SEQUENCE 180 AA; 19963 MW; 8935B0CADA923F75 CRC64;

Query Match 56.8%; Score 46; DB 2; Length 180;
Best Local Similarity 71.4%; Pred. No. 20;
Matches 10; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      2 LEMPVLVPVPPFV 15
Db      29 LECSVLVPVPPFV 42

RESULT 5
Q7RZ84 ID Q7RZ84 PRELIMINARY; PRT; 935 AA.
AC Q7RZ84
DT 01-MAR-2004 (TREMBlrel. 26, Created)
DT 01-MAR-2004 (TREMBlrel. 26, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Hypothetical protein.
GN Name=NCU04389.1;
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OR74A;
RA Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,
RA Jaffe D., FitzHugh W., Ma L.-J., Smirnov S., Purcell S., Rehman B.,
RA Elkins T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M.,

```

RA Qui D., Ianakiev P., Pedersen D., Nelson M., Washburne M.,
 RA Selitrennikoff C.P., Kinsey J.A., Braun E.L., Zelter A., Schulte U.,
 RA Kothe G.O., Jedd G., Mewes W., Straben C., Marcotte E., Greenberg D.,
 RA Roy A., Foley K., Naylor J., Thomann N., Barrett R., Gnerre S.,
 RA Kamal M., Kanvayseles M., Maucelli E., Bieleke C., Rudd S., Frishman D.,
 RA Krystofova S., Rasmussen C., Metznerberg R.L., Perkins D.D., Kroken S.,
 RA Cogoni C., Macino G., Catchside D., Li W., Pratt R.J., Osmani S.A.,
 RA DeSouza C.C., Glass L., Orbach M.J., Berglund J., Voelker R.,
 RA Yarden O., Piamann M., Seiler S., Dunlap J., Radford A., Aramayo R.,
 RA Natvig D.O., Alex L.A., Mannhaupt G., Eboile D.J., Freitag M.,
 RA Paulsen I., Sachs M.S., Lander E.S., Nussbaum C., Birren B.,
 RA "The Genome Sequence of the Filamentous Fungus Neurospora crassa."
 RL Nature 0:0-0(2003).
 CC -1- SIMILARITY: Contains 1 SET domain.
 CC -1- CAUTION: The sequence shown here is derived from an
 CC ENBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 CC EMBL; AABX01000720; EAA28230.1; -.
 DR GO; GO:0003677; F:DNA binding; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro; IPR011011; FYVE_PHD_Znf.
 DR InterPro; IPR001214; Znf_PHD.
 DR InterPro; IPR001965; Znf_PHD.
 DR Pfam; PF00628; PHD; 1.
 DR Pfam; PF00856; SET; 1.
 DR PROSITE; PS50280; SET; 1.
 DR Hypothetical protein.
 KW Hypothetical protein.
 SQ SEQUENCE 935 AA; 101487 MW; FC41A43B126C4DAF CRC64;

 Query Match 56.8%; Score 46; DB 2; Length 935;
 Best Local Similarity 61.5%; Pred. No. 1.2e+02;
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

 Qy 1 DLEMPVLPVEPPF 13
 Db 809 DLRVMPVPVPPF 821

 RESULT 6
 Q6M999 PRELIMINARY; PRT; 948 AA.
 ID Q6M999;
 AC Q6M999;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Hypothetical protein 29E8.260.
 GN Name=29E8.260;
 OS Neurospora crassa.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
 OX NCBI_TaxID=5141;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Schulte U., Aign V., Hoheisel J., Brandt P., Fartmann B., Holland R.,
 RA Nyakatura G., Mewes H.W., Mannhaupt G.;
 RL Submitted (JAN-2004) to the ENBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA German Neurospora genome project;
 RL Submitted (JAN-2004) to the ENBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: Contains 1 SET domain.
 CC EMBL; BX908809; CAF06059.1; -.
 DR GO; GO:0003677; F:DNA binding; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro; IPR011011; FYVE_PHD_Znf.
 DR InterPro; IPR001214; SET.
 DR InterPro; IPR001965; Znf_PHD.
 DR Pfam; PF00628; PHD; 1.
 DR Pfam; PF00856; SET; 1.
 DR SMART; SM00249; PHD; 1.
 DR SMART; SM00317; SET; 1.
 DR PROSITE; PS50280; SET; 1.
 DR Hypothetical protein.

SQ SEQUENCE 948 AA; 102502 MW; D53CFA84C8F6351C CRC64;

 Query Match 56.8%; Score 46; DB 2; Length 948;
 Best Local Similarity 61.5%; Pred. No. 1.2e+02;
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

 Qy 1 DLEMPVLPVEPPF 13
 Db 809 DLRVMPVPVPPF 821

 RESULT 7
 Q99ME4 PRELIMINARY; PRT; 79 AA.
 ID Q99ME4;
 AC Q99ME4;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE Thyroid hormone-response protein-1.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Xie C., Yang Y., Yang Y., Cai D., Cheng G., Li G., Luo M.,
 RL Submitted (FEB-2001) to the ENBL/GenBank/DBJ databases.
 DR EMBL; AF348365; AAK15769.1; -.
 SQ SEQUENCE 79 AA; 8994 MW; F884278152833C09 CRC64;

 Query Match 55.6%; Score 45; DB 2; Length 79;
 Best Local Similarity 63.6%; Pred. No. 12;
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

 Qy 3 EMPVLPVEPPF 13
 Db 8 EVPVLPQLPLP 18

 RESULT 8
 Q8LF16 PRELIMINARY; PRT; 80 AA.
 ID Q8LF16;
 AC Q8LF16;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE Hypothetical protein.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Haas B.J., Volkovsky N., Town C.D., Troukhan M., Alexandrov N.,
 RA Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;
 RT "Full-length messenger RNA sequences greatly improve genome
 RL annotation."
 RL Genome Biol. 3:RESEARCH0029-RESEARCH0029(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
 RA Feldmann K.;
 RL Submitted (MAR-2002) to the ENBL/GenBank/DBJ databases.
 DR EMBL; AY085097; AAM61651.1; -.
 KW Hypothetical protein.
 SQ SEQUENCE 80 AA; 8671 MW; BB1EF444B1A34E81 CRC64;

 Query Match 55.6%; Score 45; DB 2; Length 80;
 Best Local Similarity 77.8%; Pred. No. 12;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Qy 5 PVLVPVEPP 13
Db 38 PVIPTPEPP 46

RESULT 9
Q8S8R5 Q8S8R5 PRELIMINARY; PRT; 91 AA.
AC Q8S8R5;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Expressed protein.
GN Name=At2G02515;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Lin X., Kaul S., Town C.D., Benito M.-I., Cressy T.H., Haas B.J.,
RA Wu D., Maiti R., Ronning C.M., Koo H., Fujii C.Y., Utterback T.R.,
RA Barnstead M.E., Bowman C.L., White O., Nierman W.C., Fraser C.M.;
RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RL [2]
RP SEQUENCE FROM N.A.
RA Town C.D., Kaul S.;
RA Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
RL EMBL: AC004136; ANM14912.1;
SQ SEQUENCE 91 AA; 9870 MW; CF45DSB3FB66192D CRC64;

Query Match 55.6%; Score 45; DB 2; Length 91;
Best Local Similarity 77.8%; Pred. No. 14;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 5 PVLVPVEPP 13
Db 49 PVIPTPEPP 57

RESULT 10
Q6GMA4 Q6GMA4 PRELIMINARY; PRT; 200 AA.
AC Q6GMA4;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE MGC81952 protein.
GN Name=MGC81952;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RA Tissue=Kidney;
RC MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Wozney D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettner M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
```

RT "Analysis of the mouse transcriptome based on functional annotation of the KINEN genome exploration research group phase 1 & 11 team,"

Search completed: April 12, 2005, 08:05:44
Job time : 73.1818 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 12, 2005, 07:49:31 ; Search time 61.6 Seconds
(without alignments)
69.064 Million cell updates/sec

Title: US-10-691-330-6
Perfect score: 62
Sequence: 1 MPQNFYKLPQM 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	62	100.0	11	4	AAB72505
2	62	100.0	11	4	AAB59311
3	62	100.0	11	4	AAB72251
4	62	100.0	11	4	AAB72537
5	62	100.0	11	5	AAO14582
6	62	100.0	11	5	AAM51041
7	62	100.0	11	5	AAE20233
8	62	100.0	11	8	ADN60300
9	62	100.0	11	8	ADS74387
10	62	100.0	12	4	AAB59342
11	40	64.5	164	8	ABO58624
12	40	64.5	511	7	ADM04386
13	40	64.5	673	6	ABP71874
14	40	64.5	716	4	ABG24050
15	40	64.5	716	7	ADC31679
16	40	64.5	716	7	ADC33261
17	39	62.9	582	2	AAI37674
18	39	62.9	582	2	AAW97215
19	39	62.9	582	6	ABU27392
20	39	62.9	591	5	ABB49281
21	39	62.9	591	6	ABU32645
22	39	62.9	3427	7	ADJ69818
23	39	62.9	3460	5	ABB05007
24	39	62.9	3460	7	ADE60143
25	39	62.9	3460	7	ADN95182

26	39	62.9	3460	8	ADQ20919	Human sof
27	39	62.9	3460	8	ADQ88204	Human 353
28	39	62.9	3461	5	ABB05008	Mouse ree
29	39	62.9	3461	5	ABB57065	Mouse isc
30	39	62.9	3461	7	ADE60142	Rat Prote
31	39	62.9	3470	4	ABG25297	Novel hum
32	38	61.3	287	8	ADO61921	Transcrip
33	37	59.7	114	4	ABB67449	Drosophil
34	37	59.7	200	4	AAU19167	Human G p
35	37	59.7	218	4	AAO13220	Human pol
36	37	59.7	433	2	AAR98454	Oligodend
37	37	59.7	433	3	AAI57091	Human oli
38	37	59.7	490	6	ADA32995	Acinetoba
39	37	59.7	491	2	AAW52826	Human chr
40	37	59.7	491	8	ABM80154	Tumour-as
41	37	59.7	708	5	ADK37010	Novel hum
42	37	59.7	834	2	AAW52820	Human PRC
43	36	58.1	58	5	ABG56607	Human bre
44	36	58.1	62	5	ABP11495	Human ORF
45	36	58.1	126	2	AAR92475	C2Gnt cat

ALIGNMENTS

RESULT 1

AAB72505
ID AAB72505 standard; peptide; 11 AA.

XX AAB72505;

DT 09-MAY-2001 (first entry)

DE Colostrinin peptide #6.

KW Dermatological; oxidative stress regulator; colostrinin.

XX Unidentified.

PN W0200112650-A2.

PD 22-FEB-2001.

PF 17-AUG-2000; 2000WO-US022665.

PR 17-AUG-1999; 99US-0149310P.

PA (TEXA) UNIV TEXAS SYSTEM.

XX Stanton GJ, Hughes TK, Boldogh I;

XX WPI; 2001-218342/22.

XX Modulating oxidative stress level in a cell, involves contacting the cell with an oxidative stress regulator selected from colostrinin, its constituent peptide, analog or their combinations.

PS Claim 6; Page 25; 48pp; English.

XX The present invention relates to a method for modulating the oxidative stress level in a cell or a patient, comprising contacting the cell with, or administering to the patient, an oxidative stress regulator selected from colostrinin, or its constituent peptide (e.g. the present peptide), to change the level of an oxidising species in the cell. The method can be used to treat oxidative damage to skin, by decreasing or preventing an increase in the level of damage to a biomolecule of the patient

XX Sequence 11 AA;

Query Match 100.0%; Score 62; DB 4; Length 11;

Best Local Similarity 100.0%; Pred. No. 0.00033;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Colostrinin; immune response; cytokine; blood cell proliferation; central nervous system disorder; neurological disorder; mental disorder;

CC The present sequence represents a synthetic constituent peptide of
CC colostrinin, which can be used as a modulator in the present invention.
XX
SQ Sequence 11 AA;

Query Match 100.0%; Score 62; DB 8; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.00033;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MPQNFYKLPQM 11
| | | | | | | | | |
Db 1 MPQNFYKLPQM 11

RESULT 9

ADS74387
ID ADS74387 standard; peptide; 11 AA.

XX AC

ADS74387;

XX DT 16-DEC-2004 (first entry)

XX DE Ovine colostrinin peptide.

XX KW Colostrum; colostrinin; sheep; peptide purification.

XX OS Ovis aries.

XX FN WO2004081038-A1.

XX PD 23-SEP-2004.

XX PF 10-MAR-2004; 2004WO-GB001014.

XX PR 11-MAR-2003; 2003GB-00005552.

XX PR 08-MAR-2004; 2004GB-00005190.

XX PA (REG-) REGEN THERAPEUTICS PLC.

XX PI Georgiades JA, Polanowski A, Wilusz T, Kruzel ML;

XX DR WPI; 2004-677519/66.

XX PT Recovering peptides such as colostrinin from mammalian colostrum, by
XX mixing colostrum with alcohol to form alcohol phase containing peptides
XX and precipitate, separating alcohol phase from precipitate, and
XX PT recovering alcohol phase.

XX PS Disclosure; SEQ ID NO 2; 41pp; English.

XX CC The present sequence is that of a peptide that can be recovered from
XX CC ovine colostrinin using the method of the invention. The invention
XX CC provides a method for the recovery of peptides (especially colostrinin)
XX CC from colostrum in substantially pure, biologically active form and in
XX CC high yield. The method involves mixing the colostrum with an alcohol to
XX CC form an alcohol phase containing the colostrinin and a precipitate
XX CC containing higher molecular weight caseins and other proteins. Best
XX CC results are obtained using methanol or ethanol of at least 80%, and
XX CC preferably up to 100%, purity. The alcohol phase is then separated from
XX CC the precipitate, and the colostrinin is separated from the alcohol,
XX CC preferably by evaporation, to form a colostrinin-rich phase, which is
XX CC recovered. A precipitation agent, such as ammonium sulfate, may be added
XX CC either to the alcohol phase or, preferably, to the colostrinin-rich phase
XX CC to induce precipitation of the colostrinin peptides. The method is
XX CC generally applicable to the separation of peptides from fluids containing
XX CC higher molecular weight proteins, lipids, carbohydrates and/or nucleic
XX CC acids.

XX SQ Sequence 11 AA;

Query Match 100.0%; Score 62; DB 8; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.00033;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MPQNFYKLPQM 11
| | | | | | | | | |
Db 1 MPQNFYKLPQM 11

RESULT 10

AAB59342
ID AAB59342 standard; peptide; 12 AA.

XX AC

AAB59342;

XX DT 21-MAR-2001 (first entry)

XX DE Ewe colostrinin peptide fragment derived sequence #2.

XX KW Sheep; colostrinin; proline rich polypeptide; colostrum; immune disorder;
XX central nervous system disorder; dietary supplement; beta-amyloid plaque.

XX OS Ovis sp.

XX FN WO2000075173-A2.

XX PD 14-DEC-2000.

XX PF 02-JUN-2000; 2000WO-GB002128.

XX PR 02-JUN-1999; 99GB-00012852.

XX PA (REG-) REGEN THERAPEUTICS PLC.

XX PI Georgiades JA;

XX DR WPI; 2001-071058/08.

XX PT Peptides having an N-terminal amino acid sequence isolated from
XX colostrinin for treating e.g. disorders of the central nervous system and
XX immune system, viral and bacterial infections, and diseases characterized
XX by amyloid plaques.

XX PS Claim 8; Page 27; 63pp; English.

XX CC The present invention provides the sequences of a number of peptides
XX CC found in ewe's colostrinin. Colostrinin is the proline-rich polypeptide
XX CC fragment of colostrum. These peptides can be used in the treatment of
XX CC central nervous system disorders such as senile dementia, Parkinson's
XX CC disease, Alzheimer's disease, psychosis and neurosis, immune system
XX CC disorders such as bacterial and viral infections, to improve the
XX CC development of a child's immune system, as a dietary supplement, and to
XX CC promote the dissolution of beta-amyloid plaques

XX SQ Sequence 12 AA;

Query Match 100.0%; Score 62; DB 4; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.00036;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MPQNFYKLPQM 11
| | | | | | | | | |
Db 2 MPQNFYKLPQM 12

RESULT 11

ABO58624

ID ABO58624 standard; protein; 164 AA.

XX AC

ABO58624;

XX DT 29-JUL-2004 (first entry)

XX DE Human genome derived single exon protein #4858.

XX KW Human; gene expression; single exon probe; microarray;

XX Unidentified.
 OS CNI364813-A.
 PN 21-AUG-2002.
 PD 10-JAN-2001; 2001CN-00105152.
 PP 10-JAN-2001; 2001CN-00105152.
 PR (BIOW-) BIOWINDOW GENE DEV INC SHANGHAI.
 XX Mao Y, Xie Y;
 PI WPI; 2003-000524/01.
 XX N-PSDB; ABZ75217.
 DR New polypeptide-growth hormone 74.03 and polynucleotide for encoding such
 PT polypeptide.
 XX Claim 1; Page 27-28 (Disclosure); 34pp; Chinese.
 PS The invention relates to the novel polypeptide, growth hormone 74.03, and
 CC the polynucleotide encoding it. The polypeptide is useful in treating
 CC various diseases, such as cardiovascular diseases, development disorder,
 CC pregnancy relative disease, various tumors, immunological diseases and
 CC some inflammations. The invention also discloses the antagonist resisting
 CC the polypeptide and its treatment effect, and the application of the
 CC polynucleotide. The present sequence represents the growth hormone 74.03
 CC of the invention
 XX Sequence 673 AA;
 SQ Query Match 64.5%; Score 40; DB 6; Length 673;
 Best Local Similarity 60.0%; Pred. No. 1.7e+02;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 MPQNFYKLPQ 10
 Db 353 LPKSFYKAPQ 362
 :|:|:|:|:|
 RESULT 14
 ID ABG24050 standard; protein; 716 AA.
 XX AC ABG24050;
 XX DT 18-FEB-2002 (first entry)
 XX DE Novel human diagnostic protein #24041.
 XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX OS Homo sapiens.
 XX PN WO200175067-A2.
 XX PD 11-OCT-2001.
 XX PF 30-MAR-2001; 2001WO-US008631.
 XX PR 31-MAR-2000; 2000US-00540217.
 XX PR 23-AUG-2000; 2000US-00649167.
 XX PA (HYSE-) HYSEQ INC.
 XX PI Drmanac RT, Liu C, Tang YT;
 XX WPI; 2001-639362/73.
 DR N-PSDB; AAS88237.

XX -New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 XX Claim 20; SEQ ID NO 54409; 103pp; English.
 XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activities. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic
 CC amino acid sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX Sequence 716 AA;
 SQ Query Match 64.5%; Score 40; DB 4; Length 716;
 Best Local Similarity 60.0%; Pred. No. 1.8e+02;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 MPQNFYKLPQ 10
 Db 406 LPKSFYKAPQ 415
 :|:|:|:|:|
 RESULT 15
 ID ADC31679 standard; protein; 716 AA.
 XX AC ADC31679;
 XX DT 18-DEC-2003 (first entry)
 XX DE Human novel polypeptide sequence, SEQ ID NO:1761.
 XX KW Human; diagnostic; drug screening; forensics; gene mapping;
 KW biodiversity assessment; Parkinson's disease; Alzheimer's disease;
 KW neurodegenerative diseases; anaemia; platelet disorder; wound; burns;
 KW ulcers; osteoporosis; autoimmune disease; cancer;
 KW molecular weight marker; food supplement; antiparkinsonian; nootropic;
 KW neuroprotective; antianaemic; anticoagulant; thrombolytic; vulnery;
 KW antiulcer; osteopathic; immunosuppressive; antiinflammatory; cytostatic;
 KW gene therapy; chromosome 6q22.1-22.33.
 XX OS Homo sapiens.
 XX PN WO2003029271-A2.
 XX PD 10-APR-2003.
 XX PF 24-SEP-2002; 2002WO-US030474.
 XX PR 24-SEP-2001; 2001US-0324631P.
 XX PA (HYSE-) HYSEQ INC.
 XX PI Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;

Tue Apr 12 09:04:58 2005

us-10-691-330-6.rag

PI Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;
 PI Haley-Vicente D, Drmanac RT;
 XX WPI; 2003-371981/35.
 DR N-PSDB; ADC30708.
 XX
 XX New polynucleotide and polypeptide useful for diagnosing, preventing or
 PT treating conditions such as neurodegenerative diseases, anemias, platelet
 PT disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
 PT cancer.
 XX
 XX Claim 20; SEQ ID NO 1761; 1185pp; English.
 XX
 XX The invention relates to 971 novel human cDNA sequences (ADC29919-
 CC ADC30889) and the polypeptides they encode (ADC30890-ADC31860). The
 CC invention also relates to nucleic acid sequences over 99% identical with
 CC the novel human cDNAs. The invention additionally encompasses expression
 CC vectors and host cells comprising a nucleic acid of the invention; the
 CC recombinant production of a polypeptide of the invention; an antibody
 CC against a polypeptide of the invention; a method of detecting
 CC polynucleotides or polypeptides of the invention; and methods of
 CC identifying a compound which binds to a polypeptide of the invention. The
 CC invention further discloses methods of preventing, treating or
 CC ameliorating a medical condition, kits comprising polynucleotide probes
 CC and/or monoclonal antibodies for carrying out the methods of the
 CC invention; methods for the identification of compounds that modulate the
 CC expression or activity of the polynucleotide and/or polypeptide; and 767
 CC contig sequences corresponding to the cDNA sequences of the invention
 CC (ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628
 CC -ADC33394). The nucleic acids and polypeptides of the invention are
 CC useful in diagnostics, drug screening, forensics, gene mapping, in the
 CC identification of mutations responsible for genetic disorders or other
 CC traits, for assessing biodiversity, and in producing many other types of
 CC data and products dependent on DNA and amino acid sequences. They are
 CC also used for treating diseases such as Parkinson's disease, Alzheimer's
 CC disease and other neurodegenerative diseases, anaemia, platelet
 CC disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
 CC cancer. The nucleic acids may also be used as hybridisation probes or
 CC primers, and in the recombinant production of a protein. The polypeptides
 CC are also useful in generating antibodies, as molecular weight markers,
 CC and as food supplements. The present sequence represents a specifically
 CC claimed human polypeptide sequence of the invention. Note: The sequence
 CC data for this patent did not form part of the printed specification, but
 CC was obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 716 AA;
 SQ
 Query Match 64.5%; Score 40; DB 7; Length 716;
 Best Local Similarity 60.0%; Pred No. 1.8e+02;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 MPQNFYKLPQ 10
 Db :||:|||||
 406 LFKSYKAPQ 415

Search completed: April 12, 2005, 08:18:23
 Job time : 63.6 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 12, 2005, 07:49:30 ; Search time 12.9 Seconds
(without alignments)
63.654 Million cell updates/sec

Title: US-10-691-330-6
Perfect score: 62
Sequence: 1 MPQNFYKLPQM 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*

1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PTCUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	62	100.0	11	4	US-09-641-803-6
2	40	64.5	66	4	US-09-248-796A-24603
3	39	62.9	582	2	US-08-899-244-2
4	39	62.9	582	3	US-09-224-772-2
5	39	62.9	582	3	US-09-227-804-2
6	39	62.9	3460	3	US-09-334-220-1
7	39	62.9	3461	3	US-09-334-220-2
8	38	61.3	319	4	US-09-248-796A-15493
9	37	59.7	411	4	US-09-270-767-35851
10	37	59.7	411	4	US-09-270-767-51068
11	37	59.7	440	4	US-09-538-092-999
12	37	59.7	451	4	US-09-949-016-9282
13	37	59.7	490	4	US-09-328-352-4282
14	36	58.1	89	4	US-09-248-796A-22239
15	36	58.1	126	1	US-08-118-906-4
16	36	58.1	126	1	US-08-486-196-4
17	36	58.1	126	1	US-08-488-135-4
18	36	58.1	126	2	US-08-474-085-4
19	36	58.1	211	4	US-09-134-000C-5667
20	36	58.1	360	4	US-09-252-991A-23023
21	36	58.1	382	3	US-08-660-645A-7
22	36	58.1	382	3	US-09-298-718-7
23	36	58.1	382	3	US-09-546-969-7
24	36	58.1	382	3	US-08-980-832-5
25	36	58.1	382	4	US-09-547-267-7
26	36	58.1	382	4	US-09-920-923B-5
27	36	58.1	428	1	US-07-955-041-4

28	36	58.1	428	1	US-08-227-455-4
29	36	58.1	428	1	US-08-472-482-4
30	36	58.1	428	1	US-08-487-069-4
31	36	58.1	428	3	US-09-233-506-3
32	36	58.1	428	4	US-09-645-192-13
33	36	58.1	428	4	US-10-084-406-13
34	36	58.1	1038	4	US-09-081-385-151
35	35	56.5	45	4	US-09-257-179-64
36	35	56.5	464	4	US-09-543-681A-5655
37	35	56.5	542	4	US-09-589-733C-2
38	35	56.5	542	4	US-09-589-733C-20
39	35	56.5	562	4	US-09-270-767-43731
40	35	56.5	761	3	US-09-625-188-14
41	35	56.5	775	4	US-09-107-433-3888
42	35	56.5	844	4	US-09-949-016-9438
43	35	56.5	1033	4	US-09-583-110-5097
44	35	56.5	1042	3	US-09-387-695-2
45	34.5	55.6	596	3	US-08-481-190-8

ALIGNMENTS

RESULT 1
US-09-641-803-6
; Sequence 6, Application US/09641803
; Patent No. 6500798
; GENERAL INFORMATION:
; APPLICANT: STANTON, G. John
; APPLICANT: HUGHES, Thomas K.
; APPLICANT: BOLDOGH, Istvan
; TITLE OF INVENTION: USE OF COLOSTRININ, CONSTITUENT PEPTIDES THEREOF, AND ANALOGS THEREOF AS OXIDATIVE STRESS REGULATORS
; FILE REFERENCE: 265.00220101
; CURRENT APPLICATION NUMBER: US/09/641,803
; PRIOR FILING DATE: 2000-08-17
; PRIOR APPLICATION NUMBER: 60/149,310
; PRIOR FILING DATE: 1999-08-17
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 6
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: peptide
US-09-641-803-6

Query Match 100.0%; Score 62; DB 4; Length 11;
Best Local Similarity 100.0%; Pred. No. 4.5e-05;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MPQNFYKLPQM 11
| | | | | | | | | |
Db 1 MPQNFYKLPQM 11

RESULT 2
US-09-248-796A-24603
; Sequence 24603, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13

```
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 24603
; LENGTH: 66
; TYPE: PRT
; ORGANISM: Candida albicans
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (12)
; OTHER INFORMATION: Identity of amino acid sequences at the above locations are unknown
US-09-248-796A-24603

Query Match      64.5%; Score 40; DB 4; Length 66;
Best Local Similarity 60.0%; Pred. No. 2.9;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MPQNFYKLPQ 10
Db 30 LPNNFYNLPK 39

RESULT 3
US-08-899-244-2
; Sequence 2, Application US/08899244
; Patent No. 5882892
; GENERAL INFORMATION:
; APPLICANT: Reichard, Raymond W.
; APPLICANT: Brown, James R.
; APPLICANT: Lawlor, Elizabeth J.
; TITLE OF INVENTION: NOVEL asps
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dechert Price & Rhoads
; STREET: 997 Lenox Drive, Building 3, Suite 210
; CITY: Lawrenceville
; STATE: NJ
; COUNTRY: USA
; ZIP: 08543
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/899,244
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Bloom, Allen
; REGISTRATION NUMBER: 29,135
; REFERENCE/DOCKET NUMBER: GM10049
; TELEPHONE: 609-520-3214
; TELEFAX: 609-520-3259
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 582 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
US-08-899-244-2

Query Match      62.9%; Score 39; DB 2; Length 582;
Best Local Similarity 77.8%; Pred. No. 42;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 POFNYKLPQ 10
Db 190 PGNFYALPQ 198

RESULT 5
US-09-227-804-2
; Sequence 2, Application US/09227804
; Patent No. 6265188
; GENERAL INFORMATION:
; APPLICANT: Reichard, Raymond W.
; APPLICANT: Brown, James R.
; APPLICANT: Lawlor, Elizabeth J.
; TITLE OF INVENTION: NOVEL asps
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dechert Price & Rhoads
; STREET: 997 Lenox Drive, Building 3, Suite 210
; CITY: Lawrenceville
; STATE: NJ
; COUNTRY: USA
; ZIP: 08543
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/224,772
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/899,244
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Bloom, Allen
; REGISTRATION NUMBER: 29,135
; REFERENCE/DOCKET NUMBER: GM10049
; TELEPHONE: 609-520-3214
; TELEFAX: 609-520-3259
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 582 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
US-09-224-772-2

Query Match      62.9%; Score 39; DB 3; Length 582;
Best Local Similarity 77.8%; Pred. No. 42;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 POFNYKLPQ 10
Db 190 PGNFYALPQ 198
```

```
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/227,804
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/899,244
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: BLOOM, Allen
; REGISTRATION NUMBER: 29,135
; REFERENCE/DOCKET NUMBER: GM10049
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 609-520-3214
; TELEFAX: 609-520-3259
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 582 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
US-09-227-804-2

Query Match 62.9%; Score 39; DB 3; Length 582;
Best Local Similarity 77.8%; Pred. No. 42;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 MPQNFYKLPQ 10
Db 190 PQNFYALPQ 198

RESULT 6
US-09-334-220-1
; Sequence 1, Application US/09334220
; Patent No. 6323177
; GENERAL INFORMATION:
; APPLICANT: St. Jude's Children's Research Hospital
; APPLICANT: Curran, Thomas
; TITLE OF INVENTION: INTERACTION OF REELIN WITH VERY LOW
; TITLE OF INVENTION: DENSITY LIPOPROTEIN (VLDL) RECEPTOR FOR SCREENING AND
; TITLE OF INVENTION: THERAPIES
; FILE REFERENCE: 2427/0F704
; CURRENT APPLICATION NUMBER: US/09/334,220
; CURRENT FILING DATE: 1999-06-16
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 3460
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-334-220-1

Query Match 62.9%; Score 39; DB 3; Length 3460;
Best Local Similarity 66.7%; Pred. No. 2.7e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MPQNFYKLP 9
Db 1235 LPQNFYKLP 1243

RESULT 7
US-09-334-220-2
; Sequence 2, Application US/09334220
; Patent No. 6323177
; GENERAL INFORMATION:
; APPLICANT: St. Jude's Children's Research Hospital
; APPLICANT: Curran, Thomas
; TITLE OF INVENTION: INTERACTION OF REELIN WITH VERY LOW
; TITLE OF INVENTION: DENSITY LIPOPROTEIN (VLDL) RECEPTOR FOR SCREENING AND
; TITLE OF INVENTION: THERAPIES
; FILE REFERENCE: 2427/0F704
; CURRENT APPLICATION NUMBER: US/09/334,220
; CURRENT FILING DATE: 1999-06-16
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 3461
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-334-220-2

Query Match 62.9%; Score 39; DB 3; Length 3461;
Best Local Similarity 66.7%; Pred. No. 2.7e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MPQNFYKLP 9
Db 1236 LPQNFYKLP 1244

RESULT 8
US-09-248-796A-15493
; Sequence 15493, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 15493
; LENGTH: 319
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-15493

Query Match 61.3%; Score 38; DB 4; Length 319;
Best Local Similarity 66.7%; Pred. No. 34;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 PQNFYKLPQ 10
Db 82 PQNFYKLPQ 90

RESULT 9
US-09-270-767-35851
; Sequence 35851, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 35851
; LENGTH: 411
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
```

```

; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-35851

Query Match      59.7%; Score 37; DB 4; Length 411;
Best Local Similarity 87.5%; Pred. No. 68;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 MPQNFYKL 8
Db      167 MPQNFQKL 174

RESULT 10
US-09-270-767-51068
; Sequence 51068, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 51068
; LENGTH: 411
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-51068

Query Match      59.7%; Score 37; DB 4; Length 411;
Best Local Similarity 87.5%; Pred. No. 68;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 MPQNFYKL 8
Db      167 MPQNFQKL 174

RESULT 11
US-09-538-092-999
; Sequence 999, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Giot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CuraPatSeqFormer Version 0.9
; SEQ ID NO 999
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Polypeptide Accession Number P23515
US-09-538-092-999

Query Match      59.7%; Score 37; DB 4; Length 440;
Best Local Similarity 60.0%; Pred. No. 73;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy      1 MPQNFYKL 10

RESULT 12
US-09-949-016-9282
; Sequence 9282, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9282
; LENGTH: 451
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-9282

Query Match      59.7%; Score 37; DB 4; Length 451;
Best Local Similarity 60.0%; Pred. No. 75;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy      1 MPQNFYKL 10
Db      397 MPNFMSEMPQ 406

RESULT 13
US-09-328-352-4282
; Sequence 4282, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 4282
; LENGTH: 490
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-4282

Query Match      59.7%; Score 37; DB 4; Length 490;
Best Local Similarity 60.0%; Pred. No. 82;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy      2 PQNFYKL 11
Db      140 PQTLLRLPQM 149

RESULT 14
US-09-248-796A-22239
; Sequence 22239, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS

```

Search completed: April 12, 2005, 08:07:53
Job time : 13.9 secs

FILE REFERENCE: 107196.132
CURRENT APPLICATION NUMBER: US/09/248,796A
CURRENT FILING DATE: 1999-02-12
PRIOR APPLICATION NUMBER: US 60/074,725
PRIOR FILING DATE: 1998-02-13
PRIOR APPLICATION NUMBER: US 60/096,409
PRIOR FILING DATE: 1998-08-13
NUMBER OF SEQ ID NOS: 28208
SEQ ID NO 22239
LENGTH: 89
TYPE: PRT
ORGANISM: Candida albicans
US-09-248-796A-22239

Query Match 58.1%; Score 36; DB 4; Length 89;
Best Local Similarity 66.7%; Pred. No. 21;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
Qy 1 MPQNFYKLP 9
Db 56 IPSNFYKEP 64

RESULT 15
US-08-118-906-4
Sequence 4, Application US/08118906
Patent No. 5484590
GENERAL INFORMATION:
APPLICANT: Fukuda, Minoru
TITLE OF INVENTION: Expression of the Developmental I
TITLE OF INVENTION: Antigen By a Cloned Human cDNA Encoding a Member of a
TITLE OF INVENTION: Beta1,6-N-Acetylglucosaminyltransferase Gene Family
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell and Flores
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/118,906
FILING DATE: 09-SEP-1993
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LJ 9526
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 126 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-118-906-4

Query Match 58.1%; Score 36; DB 1; Length 126;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MPQNFY 6
Db 37 MPQNFY 42

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 12, 2005, 07:49:30 ; Search time 34.1 Seconds
(without alignments)
107.096 Million cell updates/sec

Title: US-10-691-330-6
Perfect score: 62
Sequence: 1 MPQNFYKLPQM 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1418010 seqs, 33197259 residues

Total number of hits satisfying chosen parameters: 1418010

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
1	62	100.0	11	US-10-281-652-6
2	62	100.0	11	US-10-691-330-6
3	40	64.5	164	US-10-029-386-32258
4	40	64.5	511	US-10-108-260A-3071
5	40	64.5	878	US-10-825-692-92
6	39	62.9	582	US-10-282-122A-55316
7	39	62.9	591	US-10-282-122A-60569
8	39	62.9	3427	US-10-408-765A-1624
9	39	62.9	3460	US-10-753-267-46
10	38	61.3	248	US-10-424-599-250800
11	38	61.3	547	US-10-767-701-45737
12	37.5	60.5	85	US-10-424-599-150738
13	37	59.7	200	US-09-801-944B-143

```

14 36 58.1 55 15 US-10-424-599-188125 Sequence 188125,
15 36 58.1 58 14 US-10-082-830-236 Sequence 236, App
16 36 58.1 73 15 US-10-424-599-247522 Sequence 247522,
17 36 58.1 94 15 US-10-424-599-199418 Sequence 199418,
18 36 58.1 157 16 US-10-767-701-54723 Sequence 54723, A
19 36 58.1 322 15 US-10-282-122A-66811 Sequence 66811, A
20 36 58.1 382 9 US-09-547-267-7 Sequence 7, Appli
21 36 58.1 382 10 US-09-920-923-5 Sequence 5, Appli
22 36 58.1 382 15 US-10-695-980-5 Sequence 5, Appli
23 36 58.1 428 9 US-09-797-207-14 Sequence 14, Appl
24 36 58.1 428 14 US-10-084-406-13 Sequence 13, Appl
25 36 58.1 428 14 US-10-388-307-13 Sequence 13, Appl
26 36 58.1 428 15 US-10-445-318-14 Sequence 14, Appl
27 36 58.1 428 17 US-10-939-066-11 Sequence 11, Appl
28 36 58.1 715 13 US-10-087-192-1998 Sequence 1998, Ap
29 36 58.1 748 13 US-10-087-192-1995 Sequence 1995, Ap
30 36 58.1 920 13 US-10-087-192-1518 Sequence 1518, Ap
31 36 58.1 954 16 US-10-437-963-166760 Sequence 166760,
32 36 58.1 1038 9 US-09-752-639-151 Sequence 151, App
33 36 58.1 1038 9 US-09-984-198-151 Sequence 151, App
34 36 58.1 1349 16 US-10-437-963-144453 Sequence 144453,
35 36 58.1 1953 15 US-10-369-493-1945 Sequence 1945, Ap
36 35 56.5 45 9 US-09-729-835-64 Sequence 64, Appl
37 35 56.5 45 15 US-10-373-809-64 Sequence 64, Appl
38 35 56.5 50 15 US-10-369-324-100 Sequence 100, App
39 35 56.5 50 16 US-10-607-538-125 Sequence 125, App
40 35 56.5 146 16 US-10-437-963-125125 Sequence 125125,
41 35 56.5 237 15 US-10-424-599-191896 Sequence 191896,
42 35 56.5 332 9 US-09-738-626-5674 Sequence 5674, Ap
43 35 56.5 429 16 US-10-437-963-190212 Sequence 190212,
44 35 56.5 542 15 US-10-636-396-2 Sequence 2, Appli
45 35 56.5 542 15 US-10-636-396-20 Sequence 20, Appl

```

ALIGNMENTS

```

RESULT 1
US-10-281-652-6
; Sequence 6, Application US/10281652
; Publication No. US20030091608A1
; GENERAL INFORMATION:
; APPLICANT: STANTON, G. John
; APPLICANT: HUGHES, Thomas K.
; APPLICANT: BOLDGOH, Istvan
; TITLE OF INVENTION: USE OF COLOSTRININ, CONSTITUENT PEPTIDES THEREOF, AND
; FILE REFERENCE: 265.00220101
; CURRENT APPLICATION NUMBER: US/10/281.652
; CURRENT FILING DATE: 2002-10-28
; PRIOR FILING DATE: 2000-08-17
; PRIOR APPLICATION NUMBER: 60/149,310
; PRIOR FILING DATE: 1999-08-17
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: peptide
US-10-281-652-6

```

Query Match 100.0%; Score 62; DB 14; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.00031;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MPQNFYKLPQM 11
Db 1 MPQNFYKLPQM 11

```

RESULT 2
US-10-691-330-6
; Sequence 6, Application US/10691330
; Publication No. US20050042300A1
; GENERAL INFORMATION:
; APPLICANT: Boldogh, Istvan
; APPLICANT: Stanton, G. John
; APPLICANT: Georgiades, Jerzy A.
; APPLICANT: Hughes, Thomas K., Jr.
; APPLICANT: Kruzel, Marian
; TITLE OF INVENTION: USE OF COLOSTRININ, CONSTITUENT PEPTIDES THEREOF, AND ANALOGS
; TITLE OF INVENTION: THEREOF AS INHIBITORS OF APOPTOSIS AND OTHER CELLULAR DAMAGE
; FILE REFERENCE: 265.00390101
; CURRENT APPLICATION NUMBER: US/10/691,330
; CURRENT FILING DATE: 2003-10-22
; PRIOR APPLICATION NUMBER: US 60/420,369
; PRIOR FILING DATE: 2002-10-22
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6
; LENGTH: 11
; TYPE: PRT
; ORGANISM: artificial
; FEATURE:
; OTHER INFORMATION: Synthetic Peptides
US-10-691-330-6
Query Match 100.0%; Score 62; DB 17; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.00031;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MPQNFYKLPQM 11
DB 1 MPQNFYKLPQM 11

RESULT 3
US-10-029-386-32258
; Sequence 32258, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEWICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 32258
; LENGTH: 164
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL132795.12
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.5
; OTHER INFORMATION: SWISSPROT HIT: Q01656, EVALUATE 5.00e-21
US-10-029-386-32258
Query Match 64.5%; Score 40; DB 14; Length 164;
Best Local Similarity 60.0%; Pred. No. 37;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 MPQNFYKLPQ 10
DB 201 LPKSFYKAPQ 46

RESULT 4
US-10-108-260A-3071
; Sequence 3071, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: NO. US20040005560A1el full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3071
; LENGTH: 511
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-108-260A-3071
Query Match 64.5%; Score 40; DB 15; Length 511;
Best Local Similarity 60.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 MPQNFYKLPQ 10
DB 201 LPKSFYKAPQ 210

RESULT 5
US-10-825-692-92
; Sequence 92, Application US/10825692
; Publication No. US2005004232A1
; GENERAL INFORMATION:
; APPLICANT: Hotez, Peter
; APPLICANT: Ashcom, James
; APPLICANT: Bdamchian, Mahnaz
; APPLICANT: Zhan, Bin
; APPLICANT: Wang, Yan
; APPLICANT: Hawdon, John
; APPLICANT: Loukas, Alexander
; APPLICANT: Williamson, Angela
; APPLICANT: Jones, Brian
; APPLICANT: Bethony, Jeffrey
; APPLICANT: Goud, Gaddam
; APPLICANT: Mendez, Susana
; APPLICANT: Botazzi, Maria E.
; TITLE OF INVENTION: Hookworm Vaccine
; FILE REFERENCE: 0374007aa
; CURRENT APPLICATION NUMBER: US/10/825,692
; CURRENT FILING DATE: 2004-04-16
; PRIOR APPLICATION NUMBER: US 60/329,533
; PRIOR FILING DATE: 2001-10-17
; PRIOR APPLICATION NUMBER: US 60/332,007
; PRIOR FILING DATE: 2001-11-23
; PRIOR APPLICATION NUMBER: US 60/375,404
; PRIOR FILING DATE: 2002-04-26
; PRIOR APPLICATION NUMBER: PCT US02/33106
; PRIOR FILING DATE: 2002-10-17
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 92
; LENGTH: 878
; TYPE: PRT
; ORGANISM: Necator americanus
US-10-825-692-92
Query Match 64.5%; Score 40; DB 17; Length 878;
Best Local Similarity 70.0%; Pred. No. 2.1e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```


Qy 1 MPQNFYKLPQ 10
||:|||||
Db 302 MPREFYVLPQ 311

RESULT 6

US-10-282-122A-55316
; Sequence 55316, Application US/10282122A
; Publication No. US20040029129A1

GENERAL INFORMATION:

; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.

; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

; FILE REFERENCE: ELITRA.034A

; CURRENT APPLICATION NUMBER: US/10/282,122A

; CURRENT FILING DATE: 2003-02-20

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/230,335

; PRIOR FILING DATE: 2000-09-06

; PRIOR APPLICATION NUMBER: 60/230,347

; PRIOR FILING DATE: 2000-09-09

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/267,636

; PRIOR FILING DATE: 2001-02-09

; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 78614

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 55316

; LENGTH: 582

; TYPE: PRT

; ORGANISM: Chlamydia trachomatis

US-10-282-122A-55316

Query Match 62.9%; Score 39; DB 15; Length 582;

Best Local Similarity 77.8%; Pred. No. 2.1e+02;

Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 PQNFYKLPQ 10
||:|||||
Db 190 PQNFYVLPQ 198

RESULT 7

US-10-282-122A-60569
; Sequence 60569, Application US/10282122A
; Publication No. US20040029129A1

GENERAL INFORMATION:

; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl

; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.

; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

; FILE REFERENCE: ELITRA.034A

; CURRENT APPLICATION NUMBER: US/10/282,122A

; CURRENT FILING DATE: 2003-02-20

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/230,335

; PRIOR FILING DATE: 2000-09-06

; PRIOR APPLICATION NUMBER: 60/230,347

; PRIOR FILING DATE: 2000-09-09

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/267,636

; PRIOR FILING DATE: 2001-02-09

; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 78614

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 60569

; LENGTH: 591

; TYPE: PRT

; ORGANISM: Listeria monocytogenes

US-10-282-122A-60569

Query Match 62.9%; Score 39; DB 15; Length 591;

Best Local Similarity 77.8%; Pred. No. 2.1e+02;

Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 PQNFYKLPQ 10
||:|||||
Db 189 PQNFYVLPQ 197

RESULT 8

US-10-408-765A-1624

; Sequence 1624, Application US/10408765A

; Publication No. US20040101874A1

GENERAL INFORMATION:

; APPLICANT: Ghosh, Soumitra S.

; APPLICANT: Fahy, Eoin D.

; APPLICANT: Zhang, Bing

; APPLICANT: Gibson, Bradford W.

; APPLICANT: Taylor, Steven W.

; APPLICANT: Glenn, Gary M.

; APPLICANT: Warlock, Dale E.

; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION

; FILE REFERENCE: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME

; FILE REFERENCE: 660088.465

; CURRENT APPLICATION NUMBER: US/10/408,765A

; CURRENT FILING DATE: 2003-04-04

; NUMBER OF SEQ ID NOS: 3077

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1624

; LENGTH: 3427

; TYPE: PRT

```
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: VARIANT
/ LOCATION: 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124,
/ LOCATION: 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136,
/ LOCATION: 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148,
/ LOCATION: 149, 150, 151, 152, 153, 154, 155, 156, 157, 158
/ OTHER INFORMATION: Xaa = Any Amino Acid
US-10-408-765A-1624

Query Match      62.9%; Score 39; DB 16; Length 3427;
Best Local Similarity 66.7%; Pred. No. 1.3e+03;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy      1 MPQNFYKLP 9
Db      1235 LPQNFYEKP 1243

RESULT 9
US-10-753-267-46
; Sequence 46, Application US/10753267
; Publication No. US20050037946A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Stagliano, Nancy E.
; APPLICANT: Healy, Aileen
; APPLICANT: Acton, Susan L.
; APPLICANT: Galvin, Katherine M.
; APPLICANT: Donoghue, Mary A.
; APPLICANT: Rodriguez-Way, Amelie
; APPLICANT: Tomlinson, James E.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
; TITLE OF INVENTION: CARDIOVASCULAR DISEASE USING 1722, 10280, 59917, 85553,
; TITLE OF INVENTION: 10653, 12282, 21668, 17794, 2210, 6169, 10102, 21061, 17662,
; TITLE OF INVENTION: 1468, 12282, 6350, 9035, 1820, 23652, 7301, 8925, 8701,
; TITLE OF INVENTION: 3538, 9462, 9123, 12788, 17729, 65552, 1261, 21476, 33770,
; TITLE OF INVENTION: 9380, 2569654, 33556, 53656, 44143, 32612, 10671, 261,
; TITLE OF INVENTION: 44570, 43962, 2552, 2417, 19339, 43969, 8921, 8993, 955,
; TITLE OF INVENTION: 32345, 966, 1920, 17318, 1510, 14180, 26005, 554, 16408,
; TITLE OF INVENTION: 42028, 112091, 13886, 13942, 1673, 54946 OR 2419
; FILE REFERENCE: MPI03-003P1RNONNM
; CURRENT APPLICATION NUMBER: US/10/753.267
; CURRENT FILING DATE: 2004-01-08
; PRIOR APPLICATION NUMBER: US 60/439,683
; PRIOR FILING DATE: 2003-01-13
; PRIOR APPLICATION NUMBER: US 60/445,216
; PRIOR FILING DATE: 2003-02-05
; PRIOR APPLICATION NUMBER: US 60/448,036
; PRIOR FILING DATE: 2003-02-18
; PRIOR APPLICATION NUMBER: US 60/454,189
; PRIOR FILING DATE: 2003-03-12
; PRIOR APPLICATION NUMBER: US 60/457,541
; PRIOR FILING DATE: 2003-03-25
; PRIOR APPLICATION NUMBER: US 60/466,411
; PRIOR FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: US 60/469,041
; PRIOR FILING DATE: 2003-05-08
; PRIOR APPLICATION NUMBER: US 60/477,414
; PRIOR FILING DATE: 2003-06-10
; PRIOR APPLICATION NUMBER: US 60/478,560
; PRIOR FILING DATE: 2003-06-13
; PRIOR APPLICATION NUMBER: US 60/489,772
; PRIOR FILING DATE: 2003-07-24
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 130
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 46
; LENGTH: 3460
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-753-267-46
```

```
Query Match      62.9%; Score 39; DB 17; Length 3460;
Best Local Similarity 66.7%; Pred. No. 1.3e+03;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy      1 MPQNFYKLP 9
Db      1235 LPQNFYEKP 1243

RESULT 10
US-10-424-599-250800
; Sequence 250800, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285694
; SEQ ID NO 250800
; LENGTH: 248
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(248)
; OTHER INFORMATION: unsure at all xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_68501C.1.pcp
US-10-424-599-250800

Query Match      61.3%; Score 38; DB 15; Length 248;
Best Local Similarity 66.7%; Pred. No. 1.3e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy      2 PQNFYKLPQ 10
Db      172 PHNYHLPQ 180

RESULT 11
US-10-767-701-45737
; Sequence 45737, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 45737
; LENGTH: 547
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C6036_1.pcp
US-10-767-701-45737

Query Match      61.3%; Score 38; DB 16; Length 547;
Best Local Similarity 66.7%; Pred. No. 2.9e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy      2 PQNFYKLPQ 10
Db      172 PHNYHLPQ 180
```

Db 225 PQNVYQVPQ 233

RESULT 12

US-10-424-599-150738
; Sequence 150738, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 150738
; LENGTH: 85
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(85)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_107140C.1.pap
US-10-424-599-150738

Query Match 60.5%; Score 37.5; DB 15; Length 85;
Best Local Similarity 80.0%; Pred. No. 51;
Matches 8; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

Qy 2 PQNVYK-LPQ 10
|:|:|:|:|
Db 27 PQNVFKELPQ 36

RESULT 13

US-09-801-944B-143
; Sequence 143, Application US/09801944B
; Publication No. US20040014169A1
; GENERAL INFORMATION:
; APPLICANT: Vogeli, Gabriel
; APPLICANT: Wood, Linda S.
; TITLE OF INVENTION: No. US20040014169A1el G Protein-Coupled Receptors
; FILE REFERENCE: 00100US1
; CURRENT APPLICATION NUMBER: US/09/801,944B
; CURRENT FILING DATE: 2003-08-01
; PRIOR APPLICATION NUMBER: 60/187,828
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/187,715
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/187,929
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/187,930
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/187,825
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/187,833
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/187,830
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/187,829
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/187,582
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/187,581
; PRIOR FILING DATE: 2000-03-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 273
; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 143
; LENGTH: 200
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-801-944B-143

Query Match 59.7%; Score 37; DB 11; Length 200;
Best Local Similarity 75.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 3 QNVFKLPQ 10
|:|:|:|:|
Db 177 QNVFVFPQ 184

RESULT 14

US-10-424-599-188125
; Sequence 188125, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 188125
; LENGTH: 55
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_140891C.1.pap
US-10-424-599-188125

Query Match 58.1%; Score 36; DB 15; Length 55;
Best Local Similarity 62.5%; Pred. No. 60;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 2 PQNVFKLP 9
|:|:|:|:|
Db 38 PQNVFKIP 45

RESULT 15

US-10-082-830-236
; Sequence 236, Application US/10082830
; Publication No. US20030077604A1
; GENERAL INFORMATION:
; APPLICANT: Sun, Yongming
; APPLICANT: Recipon, Herve
; APPLICANT: Salceda, Susana
; APPLICANT: Liu, Chenghua
; APPLICANT: Turner, Leah
; TITLE OF INVENTION: Compositions and Methods Relating to Breast Specific
; FILE REFERENCE: DEX-0249
; CURRENT APPLICATION NUMBER: US/10/082,830
; CURRENT FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 60/243,802
; PRIOR FILING DATE: 2000-10-27
; NUMBER OF SEQ ID NOS: 282
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 236
; LENGTH: 58
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-082-830-236

Query Match 58.1%; Score 36; DB 14; Length 58;

us-10-691-330-6.rapb

Tue Apr 12 09:04:59 2005

Best Local Similarity 54.5%; Pred. No. 63;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MPQNFYKLPOM 11
Db 28 IPOSFFLLPAM 38

Search completed: April 12, 2005, 07:55:21
Job time : 34.1 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 12, 2005, 07:49:31 ; Search time 7.9 Seconds
(without alignments)
133.973 Million cell updates/sec

Title: US-10-691-330-6
Perfect score: 62
Sequence: 1 MPQNFYKLPQM 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	41	66.1	113	2 B72629	hypothetical prote
2	39	62.9	247	2 H70432	ABC transporter -
3	39	62.9	582	2 E81660	aspartyl-trNA synt
4	39	62.9	582	2 G71500	aspartate-trNA lig
5	39	62.9	591	2 AG1264	aspartyl-trNA synt
6	39	62.9	591	2 A11626	aspartyl-trNA synt
7	39	62.9	3461	2 S58870	reelin precursor -
8	38	61.3	289	2 C96752	probable DNA-bindi
9	38	61.3	445	2 F97855	hypothetical prote
10	38	61.3	445	2 B71642	hypothetical prote
11	38	61.3	1374	2 S62524	probable RNA helic
12	37	59.7	241	2 JH0365	Balbiani ring a 67
13	37	59.7	266	2 S75992	hypothetical prote
14	37	59.7	440	2 F84955	UDP-N-acetylmuromo
15	37	59.7	440	2 A39613	oligodendrocyte-my
16	37	59.7	609	2 G87483	aspartyl-trNA synt
17	37	59.7	660	1 G6BP74	baseplate protein
18	37	59.7	779	2 T21021	hypothetical prote
19	36	58.1	184	2 S26046	sex-determining pr
20	36	58.1	322	2 C82991	probable aromatic
21	36	58.1	428	2 A46293	beta-1,3-galactosy
22	36	58.1	478	2 T23792	hypothetical prote
23	36	58.1	486	2 T21481	hypothetical prote
24	36	58.1	564	2 A86145	hypothetical prote
25	36	58.1	606	2 E87318	hypothetical prote
26	36	58.1	1186	2 T03180	tyrosine protein k
27	36	58.1	1953	2 S63244	BNI1 protein - yea
28	35	56.5	182	2 S51802	vomeroneasal secret
29	35	56.5	195	2 B70151	endonuclease precu

30	35	56.5	210	2 T04316	heat shock protein
31	35	56.5	253	2 E83107	conserved hypothet
32	35	56.5	349	2 AH2858	glycosyltransferas
33	35	56.5	363	2 F97835	alpha-D-mannose-al
34	35	56.5	442	2 T02820	hypothetical prote
35	35	56.5	566	2 T46219	hypothetical prote
36	35	56.5	693	2 AF2275	cellulose synthase
37	35	56.5	1042	2 E95103	DNA polymerase III
38	35	56.5	1042	2 C97971	DNA-directed DNA p
39	35	56.5	1070	2 T34385	hypothetical prote
40	35	56.5	1072	2 T50949	verprolin related
41	34.5	55.6	527	2 S46088	hypothetical prote
42	34.5	55.6	527	2 S64060	probable membrane
43	34.5	55.6	596	1 S33540	catechol oxidase (
44	34.5	55.6	599	2 T07097	catechol oxidase (
45	34	54.8	62	2 AI2652	hypothetical prote

ALIGNMENTS

RESULT 1

B72629
hypothetical protein APE1490 - Aeropyrum pernix (strain K1)
C:Species: Aeropyrum pernix
C>Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C:Accession: B72629
R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takahawa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudooh, Y.; Yamazaki, J.; K
DNA Res. 6, 83-101, 1999
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr
A:Reference number: A72450; MUID:99310339; PMID:10382966
A:Accession: B72629
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-113 <KAW>
A:Cross-references: UNIPROT:Q9YBWO; DDBJ:AP0000061; NID:G5104821; PIDN:BA80488.1; PID:d
A:Experimental source: strain K1
C:Genetics:
A:Gene: APE1490
C:Superfamily: Aeropyrum pernix hypothetical protein APE1490

Query Match 66.1%; Score 41; DB 2; Length 113;
Best Local Similarity 54.5%; Pred. No. 1.4;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MPQNFYKLPQM 11
:|::|:
Db 15 LPSNFFSLPRM 25

RESULT 2

H70432
ABC transporter - Aquifex aeolicus
C:Species: Aquifex aeolicus
C>Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 09-Jul-2004
C:Accession: H70432
R:Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Ov
V.
Nature 392, 353-358, 1998
A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A:Reference number: A70300; MUID:98196666; PMID:9537320
A:Accession: H70432
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-247 <AQF>
A:Cross-references: UNIPROT:O67490; GB:AB000744; NID:G2983891; PIDN:AAC07448.1; PID:G298
A:Experimental source: strain VF5
C:Genetics:
A:Gene: abcT10
C:Superfamily: inner membrane protein malk; ATP-binding cassette homology
C:Keywords: ATP; nucleotide binding; P-loop
P;18-215/Domain: ATP-binding cassette homology <ABC>

Tue Apr 12 09:04:59 2005

F;35-42/Region: nucleotide-binding motif A (P-loop)

```

Query Match      62.9%; Score 39; DB 2; Length 247;
Best Local Similarity 66.7%; Pred. No. 7.6;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      2 PONFYKLPQ 10
      |:|||||
Db      130 PEDFYKYPQ 138

RESULT 3
E81660
aspartyl-tRNA synthetase TC0829 [imported] - Chlamydia muridarum (strain Nigg)
C:Species: Chlamydia muridarum, Chlamydia trachomatis MoPn
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C:Accession: E81660
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,
C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,
Nucleic Acids Res. 28, 1397-1406, 2000.
A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A:Reference number: A81500; MUID:20150255; PMID:10684935
A:Accession: E81660
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-582 <TET>
A:Cross-references: UNIPROT:Q9PUK0; GB:AE002349; GB:AE002160; NID:G7190851; PIDN:AAF3962
A:Experimental source: strain Nigg (MoPn)
C:Genetics:
A:Gene: TC0829
C:Superfamily: lysine-tRNA ligase

Query Match      62.9%; Score 39; DB 2; Length 582;
Best Local Similarity 77.8%; Pred. No. 19;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      2 PONFYKLPQ 10
      |:|||||
Db      190 PGNFYALPQ 198

RESULT 4
G71500
aspartate-tRNA ligase (EC 6.1.1.12) - Chlamydia trachomatis (serotype D, strain UW3/Cx)
N:Alternate names: aspartyl-tRNA synthetase
C:Species: Chlamydia trachomatis
C:Date: 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 09-Jul-2004
C:Accession: G71500
R:Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell,
Science 282, 754-759, 1998
A:Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia tra
A:Reference number: A71570; MUID:99000809; PMID:9784136
A:Accession: G71500
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-582 <ARN>
A:Cross-references: UNIPROT:O84546; GB:AE001325; GB:AE001273; NID:G3328980; PIDN:AAC6814
A:Experimental source: serotype D, strain UW-3/Cx
C:Genetics:
A:Gene: asps
C:Function:
A:Description: activates amino acid and transfers it to specific tRNA molecule
A:Pathway: protein biosynthesis
C:Superfamily: lysine-tRNA ligase
C:Keywords: aminoacyl-tRNA synthetase; ATP; ligase; protein biosynthesis

Query Match      62.9%; Score 39; DB 2; Length 582;
Best Local Similarity 77.8%; Pred. No. 19;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      2 PONFYKLPQ 10
      |:|||||
Db      190 PGNFYALPQ 198

```

RESULT 5

```

AG1264
aspartyl-tRNA synthetase [imported] - Listeria monocytogenes (strain EGD-e)
C:Species: Listeria monocytogenes
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C:Accession: AG1264
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkak, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schluerter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A:Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AG1264
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-591 <GLA>
A:Cross-references: UNIPROT:Q8Y709; GB:NC_003210; PIDN:CAC99597.1; PID:gl6410948; GSPDB:
A:Experimental source: strain EGD-e
C:Genetics:
A:Gene: asps
C:Superfamily: lysine-tRNA ligase

```

```

Query Match      62.9%; Score 39; DB 2; Length 591;
Best Local Similarity 77.8%; Pred. No. 19;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY      2 PONFYKLPQ 10
      |:|||||
Db      189 PGNFYALPQ 197

```

RESULT 6

```

A11626
aspartyl-tRNA synthetase [imported] - Listeria innocua (strain Clip11262)
C:Species: Listeria innocua
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C:Accession: A11626
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkak, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schluerter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A:Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: A11626
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-591 <GLA>
A:Cross-references: UNIPROT:Q92BJ4; GB:AL592022; PIDN:CAC96785.1; PID:gl6414041; GSPDB:
A:Experimental source: strain Clip11262
C:Genetics:
A:Gene: asps
C:Superfamily: lysine-tRNA ligase

```

```

Query Match      62.9%; Score 39; DB 2; Length 591;
Best Local Similarity 77.8%; Pred. No. 19;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY      2 PONFYKLPQ 10
      |:|||||
Db      189 PGNFYALPQ 197

```

RESULT 7

```

S58870
reelin precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 06-Dec-1996 #sequence_revision 06-Dec-1996 #text_change 21-Jul-2003

```

C;Accession: S58870; S71844; I49297
R;D'Arcangelo, G.; Miao, G.G.; Chen, S.C.; Soares, H.D.; Morgan, J.I.; Curran, T.
Nature 374, 719-723, 1995
A;Title: A protein related to extracellular matrix proteins deleted in the mouse mutant
A;Reference number: I49297; MUID:95231649; PMID:7715726
A;Accession: S58870
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-3461 <DAR>
A;Cross-references: EMBL:U24703; NID:902486; PID:902487
R;D'Arcangelo, G.
submitted to the EMBL Data Library, April 1995
A;Reference number: S71844
A;Accession: S71844
A;Molecule type: mRNA
A;Residues: 1-215, 'T', 217-1905, 'S', 1907-3355, 'V', 3357-3391, 'N', 3393-3461 <DA2>
A;Cross-references: EMBL:U24703; NID:902486; PID:902487
F;1-27/Domain: signal sequence #status predicted <SIG>
F;28-3461/Product: reelin #status predicted <MAT>
F;1769-1795/Domain: EGF homology <EGF>

Query Match 62.9%; Score 39; DB 2; Length 3461;
Best Local Similarity 66.7%; Pred. No. 1.3e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MPQNFYKLP 9
Db 1236 LPQNFYKLP 1244
|||:|

RESULT 8
C96752
Probable DNA-binding protein F28P22.7 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C;Accession: C96752
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federpiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A66141; MUID:21016719; PMID:11130712
A;Accession: C96752
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-289 <STO>
A;Cross-references: UNIPROT:Q9CA18; GB:AE005173; NID:G6648154; PIDN:AAF21154.1; GSPDB:GN
C;Genetics:
A;Map position: 1

Query Match 61.3%; Score 38; DB 2; Length 289;
Best Local Similarity 85.7%; Pred. No. 14;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 QNFYKLP 9
Db 186 QNFYKLP 192
|||:|

RESULT 9
F97855
Hypothetical protein RC1246 [imported] - Rickettsia conorii (strain Malish 7)
C;Species: Rickettsia conorii
C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
C;Accession: F97855
R;Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.; Ro
Science 293, 2093-2098, 2001

A;Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.
A;Reference number: A97700; MUID:21442074; PMID:11557893
A;Accession: F97855
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-445 <KUR>
A;Cross-references: UNIPROT:Q9ZG77; GB:AE006914; PIDN:AAL03784.1; PID:G15620380; GSPDB:G
C;Genetics:
A;Gene: RC1246
C;Superfamily: conserved hypothetical protein b0835

Query Match 61.3%; Score 38; DB 2; Length 445;
Best Local Similarity 50.0%; Pred. No. 22;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 2 PQNFYKLPQM 11
Db 109 PQSYNLP 118
|||:|

RESULT 10
B71642
Hypothetical protein RP808 - Rickettsia prowazekii
C;Species: Rickettsia prowazekii
C;Date: 21-Nov-1998 #sequence_revision 21-Nov-1998 #text_change 09-Jul-2004
C;Accession: B71642
R;Andersson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Sicheritz-Ponten, T.; Alsmark, U.
Nature 396, 133-140, 1998
A;Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.
A;Reference number: A71630; MUID:99039499; PMID:9823893
A;Accession: B71642
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-445 <AND>
A;Cross-references: UNIPROT:Q9ZCE8; GB:AJ235273; GB:AJ235269; NID:G3861237; PIDN:CAA1523
A;Experimental source: strain Madrid E
C;Genetics:
A;Gene: RP808
C;Superfamily: conserved hypothetical protein b0835

Query Match 61.3%; Score 38; DB 2; Length 445;
Best Local Similarity 50.0%; Pred. No. 22;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 2 PQNFYKLPQM 11
Db 109 PQSYNLP 118
|||:|

RESULT 11
S62524
Probable RNA helicase/ribonuclease SPAC8A4.08c - fission yeast (Schizosaccharomyces pombe
C;Species: Schizosaccharomyces pombe
C;Date: 16-May-1996 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C;Accession: T39130; T41432; T41192; S62524
R;Lye, G.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.
submitted to the EMBL Data Library, November 1995
A;Reference number: Z21830
A;Accession: T39130
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-1374 <LYE>
A;Cross-references: UNIPROT:Q09884; EMBL:Z66569; NID:G1052533; PIDN:CAA91518.1; PID:G105
R;Seeger, K.; Harris, D.; Wood, V.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, October 1998
A;Reference number: Z21993
A;Accession: T41432
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-1374 <SEE>
A;Cross-references: EMBL:AL032824; PIDN:CAB37423.1; GSPDB:GN00068; SPDB:SPCC584.10C
A;Experimental source: strain 972h-; cosmid c584
R;Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Seeger, K.; Harris, D.

submitted to the EMBL Data Library, April 1999

Submitted to the database: 22-NOV-2001

A;Reference number: Z21976

A;Accession: T41192

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 722-1374 <LVN>

A;Cross-references: EMBL:AL049662; PIDN:CAB41233.1; GSPDB:GN00068; SPDB:SPCC188.13c

C;Genetics: <LYE1>

A;Gene: SPAC8A4.08c

A;Map position: 1

C;Genetics: <SEEL>

A;Gene: SPCC584.10c

A;Map position: 3

F;32-39/Region: nucleotide-binding motif A (P-loop) #status atypical

F;141-146/Region: nucleotide-binding motif B

F;145-148/Region: DRAD/H motif #status atypical

Query Match	61.3%;	Score 38;	DB 2;	Length 1374;
Best Local Similarity	75.0%;	Pred. No. 75;		
Matches	6;	Conservative	2;	Mismatches
				0;
				Indels
				0;
				Gaps
				0;

QY 3 QNFYKLPQ 10
|||:||||:
779 QNFFKLPE 786

RESULT 12

JH0365
 Balbiani ring a 67K protein - midge (Chironomus thummi)
 C:Species: Chironomus thummi
 C:Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 16-Aug-2004
 C:Accession: JH0365
 R: Bogachev, S.S.; Bllinov, A.G.; Kolesnikov, N.N.; Scherbik, S.V.; Taranin, A.V.; Sebelev
 Gene 96, 241-247, 1990
 A:Title: A tissue-specific puff (balbiani ring a) in Chironomus thummi may contain a gen
 A:Reference number: JH0365; MUID:91099682; PMID:2269433
 A:Accession: JH0365
 A:Molecule type: DNA
 A:Residues: 1-241 <BOG>
 A:Cross-references: UNIPROT:P24243; GB:M63152; NID:g156521; PID:AAA62726.1; PID:g156522
 A:Note: the authors translated the codon GAT for residue 216 as Thr

Query Match.	59.7%;	Score 37;	DB 2;	Length 241;
Best Local Similarity	75.0%;	Pred. No. 18;		
Matches	6;	Conservative	2;	Mismatches
				0;
				Indels
				0;
				Gaps
				0;

QY	3	QNFYKLPQ	10
		: : : :	
Db	65	KNFFKLPQ	72

RESULT 13

S75992
Hypothetical protein - *Synechocystis* sp. (strain PCC 6803)
C.Species: *Synechocystis* sp.
A.Variety: PCC 6803
C.Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
R.Accession: S75992
R.Caneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996
A.Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis*
S.
A.Reference number: S74322; MUID: 97061201; PMID: 8905231
A.Accession: S75992
A.Status: preliminary
A.Molecule type: DNA
A.Residues: 1-266 <KAN>
A.Cross-references: UNIPROT:Q55490; EMBL:D64006; GB:AB001339; NID:G1001291; PIDN:BAAL053
A.Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
A.Superfamily: short-chain ATP-binding cassette proteins; ATP-binding cassette homology
C.Keywords: Atp; nucleotide binding; P-loop
C.Keywords: Atp-binding cassette homology <ABC>

2 54 71 / position: nucleotide-binding motif A (p-loop)

Query Match	59.7%	Score 37;	DB 2;	Length 266;
Best Local Similarity	77.8%	Pred. No. 20;		
Matches	7;	Conservative	0;	Mismatches 2;
			Indels	0;
			Gaps	0;

Qy 3 QNFYKLPQM 11
db 116 OOFYLLPQM 124

RESULT 14

F84955
UDP-N-acetylmuramoylalanine-D-glutamate ligase (EC 6.3.2.9) [imported] - Buchnera sp. (st
C:Species: Buchnera sp.
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 03-Jun-2002
C:Accession: F84955
R:Shigenobu, S.; Watanabe, H.; Hattori, M.; Sakaki, Y.; Ishikawa, H.
Nature 407, 81-86, 2000
Article: Genome sequence of the endocellular bacterial symbiont of aphids Buchnera sp. AI
WRR-20445173. PMID:10993077

A;Accession: F84955
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-440 <STO>
A;Cross-references: GB:AP000398;
A;Experimental source: strain APS

C;Genetics:
A;Gene: murD; BU218
C;Superfamily: UDP-N-acetylmuramate-alanine ligase
C;Keywords: ligase

Query Match 59.7%; Score 37; DB 2; Length 440;
Best Local Similarity 66.7%; Pred. No. 34;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0

Qy 2 PQNFYKLPQ 10
| | | | |
Db 42 PSNFIKLPQ 50

RESULT 15

A39613
oligodendrocyte-myelin glycoprotein precursor - human
C:Species: Homo sapiens (man)
C:Date: 13-Sep-1991 #sequence_revision 03-Oct-1995 #text_change 09-Jul-2004
C:Accession: A36688; A34210; A39613; A30187
R:Mikol, D.D.; Alexakos, M.J.; Bayley, C.A.; Lemons, R.S.; Le Beau, M.M.; Stefansson, K.
J. Cell Biol. 111, 2673-2679, 1990
A:Title: Structure and chromosomal localization of the gene for the oligodendrocyte-myelin
A:Reference number: A36688; MUID:91115958; PMID:2277079
A:Accession: A36688
A:Molecule type: DNA
A:Residues: 1-440 <M1>
A:Cross-references: UNIPROT:P23515; GB:X57435; NID:g23105; PIDN:CAA40684.1; PID:g23106
R:Mikol, D.D.; Gulcher, J.R.; Stefansson, K.
J. Cell Biol. 110, 471-479, 1990
A:Title: The oligodendrocyte-myelin glycoprotein belongs to a distinct family of protein
A:Accession: A324910; MUID:90130636; PMID:1688857

A;Accession: A34210

A;Molecule type: DNA
A;Residues: 8-20, 'X', 22-440 <M2>
A;Cross-references: GB:X51694; NID:G35144; PID:CAA5991.1; PID:G35145
F;Viskochil, D.; Cawthon, R.; O'Connell, P.; Xu, G.; Stevens, J.; Cult
Mol. Cell. Biol. 11, 906-912, 1991
A;Title: The gene encoding the oligodendrocyte-myelin glycoprotein is

A;Accession: A39613

A:Molecule type: mRNA
A:Residues: 1-20, 'X', 22-440 <VIS>
A:Cross-references: GB:M3623; NID:G189385; PIDN:AAA59970.1; PID:G189386
R:Mikol, D.D.; Stefansson, K.
J. Cell Biol. 106, 1273-1279, 1988

A:Title: A phosphatidylinositol-linked peanut agglutinin-binding glycoprotein in central
A:Reference number: A30187; MUID:88198371; PMID:3283151
A:Accession: A30187
A:Molecule type: protein
A:Residues: 25-29,'H','31','H','33-44','D','46-50','P','52 <MI3>
C:Genetics:
A:Gene: GDB:OMG; OMGP
A:Cross-references: GDB:127563; OMIM:164345
A:Map position: 17q11.2-17q11.2
A:Introns: #status absent
A:Note: this gene lies within an intron of GDB:NFI on the opposite strand
C:Keywords: blocked carboxyl end; glycoprotein; lipoprotein; phosphatidylinositol linkage
F:1-24/Domain: signal sequence #status predicted <SIG>
F:56-78/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>
F:79-99/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>
F:100-123/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>
F:124-146/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>
F:147-167/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>
F:168-191/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>
F:192-215/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>
F:216-227/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR8>
F:426-440/Domain: carboxyl-terminal propeptide #status predicted <CTP>
F:425/Modified site: GPI-anchor ethanolamine amidated carboxyl end (Asn) (in mature form)

Query Match 59.7%; Score 37; DB 2; Length 440;
Best Local Similarity 60.0%; Pred. No. 34;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Oy 1 MPQNFYKLPQ 10
|||:::
Db 386 MPNNFSEMPQ 395

Search completed: April 12, 2005, 07:56:54
Job time : 8.9 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 12, 2005, 07:49:31 ; Search time 52.2 Seconds
(without alignments)
107.909 Million cell updates/sec

Title: US-10-691-330-6
Perfect score: 62
Sequence: 1 MPQNFYKLPQM 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	42	67.7	465	2	Q81423 plasmodium
2	42	67.7	474	2	Q81423 plasmodium
3	41	66.1	113	2	Q81423 plasmodium
4	41	66.1	592	1	SYD_RHOBA
5	40	64.5	141	2	Q27939
6	40	64.5	262	2	Q6C1Q0
7	40	64.5	487	2	Q9NQ10
8	40	64.5	689	2	Q9Y114
9	40	64.5	837	2	Q76751
10	40	64.5	902	2	Q6UEA5
11	39	62.9	247	2	Q67490
12	39	62.9	582	1	SYD_CHLMU
13	39	62.9	582	1	SYD_CHLTR
14	39	62.9	584	1	SYD_BUCAP
15	39	62.9	586	2	Q5YR88
16	39	62.9	591	1	SYD_LISIN
17	39	62.9	591	1	SYD_LISMO
18	39	62.9	591	1	SYD_ZF1
19	39	62.9	595	1	SYD_RHIME
20	39	62.9	599	2	Q6ME91
21	39	62.9	747	2	Q64FW1
22	39	62.9	1131	2	Q853A4
23	39	62.9	1804	2	Q8WPD7
24	39	62.9	3008	2	Q6Q144
25	39	62.9	3209	1	RELN_CHICK
26	39	62.9	3460	1	RELN_HUMAN
27	39	62.9	3461	1	RELN_MOUSE
28	39	62.9	3462	1	RELN_RAT
29	38	61.3	104	2	Q84TF7
30	38	61.3	266	2	Q6TV52
31	38	61.3	289	2	Q9CA18

32 38 61.3 445 1 Y808_RICPR Q9zce8 rickettsia
33 38 61.3 445 2 Q7P9V8 Q7p9v8 rickettsia
34 38 61.3 445 2 Q92G77 Q92g77 rickettsia
35 38 61.3 445 2 Q68VU1 Q68vul rickettsia
36 38 61.3 504 2 Q86AH3 Q86ah3 dictyosteli
37 38 61.3 1374 1 DCRI_SCHPO Q09884 schistosacch
38 37 59.7 202 2 Q7WSW8 Q7wsw8 bordetella
39 37 59.7 202 2 Q89QF7 Q89qf7 bradyrhizob
40 37 59.7 221 2 Q9VH11 Q9vh11 drosophila
41 37 59.7 266 2 Q55490 Q55490 synecocyst
42 37 59.7 290 2 Q23769 Q23769 chironomus
43 37 59.7 321 2 Q6BH28 Q6bhs2 debaryomyce
44 37 59.7 440 1 MURD_BUCAI P57313 buchnera ap
45

ALIGNMENTS

RESULT 1
Q81423 PRELIMINARY; PRT; 465 AA.
AC Q81423; 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN ORFNames=PFL2100W;
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22255705; Pubmed=12358864; DOI=10.1038/nature01097;
RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
RA Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA Chan M.S., Nene V., Shallow J., Suh B., Peterson J., Angiuoli S.,
RA Perteau M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,
RA Martin D.M., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA Fraser C.M., Barrell B.;
RA "Genome sequence of the human malaria parasite Plasmodium
falciparum".
RL Nature 419:498-511(2002).
RN [2]
RP SEQUENCE FROM N.A.
RA Hyman R.W., Fung E., Conway A., Kurdi O., Mao J., Miranda M.,
RA Nakao B., Rowley D., Tamaki T., Wang F., Davis R.W.;
RA Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the ubiquitin-conjugating enzyme family.
DR EMBL; AE014851; AA036504.1; -;
DR GO; GO:0016874; F:ligase activity; IEA.
DR GO; GO:0004840; F:ubiquitin conjugating enzyme activity; IEA.
DR GO; GO:00006512; P:ubiquitin cycle; IEA.
DR InterPro; IPR000608; UBQ_conjugat.
DR Pfam; PF00179; UBQ_con; 1.
DR PRODOM; PD000461; UBQ_conjugat; 1.
DR SMART; SM00212; UBCC; 1.
DR PROSITE; PS0127; UBIQUITIN_CONJUGAT_2; 1.
KW Hypothetical protein.
SQ SEQUENCE 465 AA; 53968 MW; 60C4D3C0DDFD6E260 CRC64;

Query Match 67.7%; Score 42; DB 2; Length 465;
Best Local Similarity 63.6%; Pred. No. 32;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MPQNFYKLPQM 11

Db 114 IPRNFYCLPQI 124

```

RESULT 2
Q8R8C0 PRELIMINARY; PRT; 474 AA.
ID Q8R8C0
AC Q8R8C0;
DT 01-JUN-2002 (TReMBLrel. 21, Created)
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE S-layer homology domain.
DE OrderedLocusNames=TT2087;
OS Thermoanaerobacter tengcongensis.
OC Bacteria; Firmicutes; Clostridia; Thermoanaerobacteriales;
OC Thermoanaerobacteraceae; Thermoanaerobacter.
OX NCBI_TaxID=119072;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MB4;
RX MEDLINE=21992816; PubMed=1197336; DOI=10.1101/gr.219302;
RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,
RA Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,
RA Tan H., Chen R., Wang J., Yu J., Yang H.;
RA "A complete sequence of the T. tengcongensis genome.";
RL Genome Res. 12:689-700(2002).
DR EMBL; AB013156; AAM25250.1; -.
DR InterPro; IPR001119; SLH.
DR Pfam; PF00395; SLH; 2.
KW Complete proteome.
SQ SEQUENCE 474 AA; 53073 MW; B373A46E9E58B91E CRC64;

Query Match 67.7%; Score 42; DB 2; Length 474;
Best Local Similarity 77.8%; Pred. No. 33;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 MPQNFYKLP 9
DB 260 IPTNFYKLP 268

RESULT 3
Q9YBWO PRELIMINARY; PRT; 113 AA.
ID Q9YBWO
AC Q9YBWO;
DT 01-NOV-1999 (TReMBLrel. 12, Created)
DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Hypothetical protein APE1490.
DE OrderedLocusNames=APE1490;
OS Aeropyrum pernix.
OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcales;
OC Desulfurococaceae; Aeropyrum.
OX NCBI_TaxID=56636;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K1;
RX MEDLINE=99310339; PubMed=10382966;
RA Kawarabayasi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Anka A., Kosugi H.,
RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
RA Yamazaki J., Kishida N., Oguchi A., Aoki K.-I., Kubota K.,
RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RA "Complete genome sequence of an aerobic hyper-thermophilic
RT crenarchaeon, Aeropyrum pernix K1.";
RL DNA Res. 6:83-101(1999).
DR EMBL; AF000061; BAA80488.1; -.
DR PIR; B72629; B72629.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 113 AA; 12228 MW; 1945C293782213BE CRC64;

Query Match 66.1%; Score 41; DB 2; Length 113;
Best Local Similarity 54.5%; Pred. No. 11;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 MPQNFYKLPQM 11
DB 260 IPTNFYKLPQM 268

RESULT 4
SYD_RHOBA STANDARD; PRT; 592 AA.
ID SYD_RHOBA
AC Q7UFY6;
DT 05-JUL-2004 (Rel. 44, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Aspartyl-tRNA synthetase (EC 6.1.1.12) (Aspartate--tRNA ligase)
DE (AspRS);
GN Name=aspS; OrderedLocusNames=RB8253;
OS Rhodopirellula baltica.
OC Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;
OC Planctomycetaceae; Pirellula.
OX NCBI_TaxID=117;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1;
RX MEDLINE=22735913; PubMed=12835416; DOI=10.1073/pnas.1431443100;
RA Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,
RA Ludwig W., Gade D., Beck A., Borzym K., Heitmann K., Rabus R.,
RA Schlesner H., Amann R., Reinhardt R.;
RA "Complete genome sequence of the marine planctomycete Pirellula sp.
RT strain 1.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303(2003).
CC -!- CATALYTIC ACTIVITY: ATP + L-aspartate + tRNA(Asp) = AMP +
CC diphosphate + L-aspartyl-tRNA(Asp).
CC -!- SUBUNIT: Homodimer (by similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to the class-II aminoacyl-tRNA synthetase
CC family.

-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; BX294147; CAD78543.1; -.
DR HAMAP; MF 00044; -.
DR InterPro; IPR004524; Asps_bact.
DR InterPro; IPR004115; GAD.
DR InterPro; IPR004364; tRNA-synt_2.
DR InterPro; IPR004365; tRNA-anti.
DR Pfam; PF02938; GAD; 1.
DR Pfam; PF01152; tRNA-synt_2; 2.
DR Pfam; PF01336; tRNA-anti; 1.
DR TIGRfam; TIGR00459; asps_bact; 1.
DR PROSITE; PS00862; AA_TRNA_LIGASE_II; 1.
KW Aminoacyl-tRNA synthetase; ATP-binding; Complete proteome; Ligase;
KW Protein biosynthesis.
SQ SEQUENCE 592 AA; 66228 MW; 4F1334694E8EC3AB CRC64;

Query Match 66.1%; Score 41; DB 1; Length 592;
Best Local Similarity 77.8%; Pred. No. 64;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 PONFYKLPQ 10
DB 189 PSNFYALPQ 197

RESULT 5
Q27939 PRELIMINARY; PRT; 141 AA.
ID Q27939
AC Q27939;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)

```

01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE B-casein (Fragment).
OS Alces alces (moose).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Cervidae;
OC Odocolleinae; Alces.
OX NCBI_TaxID=9852;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96364219; PubMed=8752004;
RA Gatesy J., Hayashi C., Cronin M.A., Arcander P.;
RT "Evidence from milk casein genes that cetaceans are close relatives of
RT hippopotamid artiodactyls."
RL Mol. Biol. Evol. 13:954-963(1996).
DR EMBL; U53896; AB08403.1; -.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0005215; P:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR001588; Casein.
DR Pfam; PF00363; Casein; 1.
FT NON_TER 1 141
FT TER 141 141
SQ SEQUENCE 141 AA; 15763 MW; DC39F68595C13C72 CRC64;
Query Match 64.5%; Score 40; DB 2; Length 141;
Best Local Similarity 70.0%; Pred. No. 22;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
Qy 1 MPQNFYKLPQ 10
Db 22 LPQNFPLPQ 31
RESULT 6
Q6C1Q0
ID Q6C1Q0 PRELIMINARY; PRT; 262 AA.
AC Q6C1Q0;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Similar to w|NCU09014.1 Neurospora crassa NCU09014. 1 hypothetical
DE protein.
DE ORFNames=YAL10F14421g;
OS Yarrowia lipolytica CLIB99.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Dipodascaceae; Yarrowia.
OX NCBI_TaxID=284591;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CLIB99;
RG Genolevures;
RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
RA Lafontaine I., de Montigny J., Marck C., Neuveglise C., Talia E.,
RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,
RA Barnay S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,
RA Bolrame A., Boyer J., Catcolico L., Confanioli F., de Daruvar A.,
RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
RA Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA Nicaud J.M., Nikolski M., Oztas S., Ozier-Kalogeropoulos O.,
RA Pellenz S., Potier S., Richard G.F., Straub M.L., Suleau A.,
RA Swennene D., Tekala F., Wesolowski-Louvel M., Westhof E., Wirth B.,
RA Zeniou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,
RA Bouchier C., Caudron B., Scarpelli C., Gaillard C., Weissenbach J.,
RA Wincker P., Souciet J.L.;
RT "Genome evolution in yeasts."
RL Nature 430:35-44(2004).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CLIB99;
RA Genoscope;
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR382132; CAG78221.1; -.

KW Hypothetical protein.
SQ SEQUENCE 262 AA; 30374 MW; 424B8386AEAOE084 CRC64;
Query Match 64.5%; Score 40; DB 2; Length 262;
Best Local Similarity 66.7%; Pred. No. 42;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
Qy 1 MPQNFYKLP 9
Db 90 LPNFFKLP 98
RESULT 7
Q9NQ10
ID Q9NQ10 PRELIMINARY; PRT; 487 AA.
AC Q9NQ10;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE DJ41217.1 (Similar to radial spokehead protein) (Fragment).
GN Name=dJ41217.1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Matthews L.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL132795; CAC00771.1; -.
DR InterPro; IPR006802; Radial spoke.
DR Pfam; PF04712; Radial spoke; 1.
FT NON_TER 1 1
FT TER 1 1
SQ SEQUENCE 487 AA; 55796 MW; 1ADB7D60F7B5D098 CRC64;
Query Match 64.5%; Score 40; DB 2; Length 487;
Best Local Similarity 60.0%; Pred. No. 80;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
Qy 1 MPQNFYKLPQ 10
Db 177 LPKSFYKAPQ 186
RESULT 8
Q9Y114
ID Q9Y114 PRELIMINARY; PRT; 689 AA.
AC Q9Y114;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative zinc metalloproteinase (Fragment).
GN Name=MEP4;
OS Haemonchus contortus (Barber pole worm).
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;
OC Trichostrongyloidea; Haemonchidae; Haemonchidae; Haemonchus.
OX NCBI_TaxID=6289;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Moreduin;
RA Skuce P.J., Newlands G.F.J.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF132519; AAD41474.1; -.
DR MEROPS; M13.011; -.
DR GO; GO:0008237; F:metalloproteinase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR008753; Peptidase_M13_N.
DR Pfam; PF05649; Peptidase_M13_N; 1.
FT NON_TER 1 689
FT TER 689 689
SQ SEQUENCE 689 AA; 78353 MW; 1431BEC1CB74E50 CRC64;
Query Match 64.5%; Score 40; DB 2; Length 689;

Best Local Similarity 70.0%; Pred. No. 1.1e+02; Indels 0; Gaps 0;
Matches 7; Conservative 1; Mismatches 2;

OY 1 MPQNFYKLPQ 10
Db 373 MPREFYVLPQ 382

RESULT 9

O76751 ID O76751 PRELIMINARY; PRT; 837 AA.
AC O76751; 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative zinc metalloproteinase (Fragment).
GN Name=MEP3; contortus (Barber pole worm).
OS Haemonchus contortus; Nematoda; Chromadorea; Rhabditida; Strongylida;
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Haemonchidae; Haemonchus.
OC Trichostrongyloidea; Haemonchidae; Haemonchus.
OX NCBI_TaxID=6289;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Moredu; PubMed=10652118;
RA Smith W.D., Smith S.K., Pettit D., Newlands G.P.J., Skuce P.J.;
RT "Relative protective properties of three membrane glycoprotein
fractions from Haemonchus contortus.";
RL Parasite Immunol. 22:63-71(2000).
DR EMBL; AF080172; AAC31568.1; --
DR MEROPS; M13.011; --
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0008237; F:metalloproteinase activity; IEA.
DR GO; GO:0004245; F:neprilysin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR000718; Peptidase_M13.
DR InterPro; IPR008753; Peptidase_M13_N.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR InterPro; IPR000719; Prot_Kinase.
DR Pfam; PF01431; Peptidase_M13; 1.
DR Pfam; PF05649; Peptidase_M13_N; 1.
DR PRINTS; PR00786; NEPRILYSIN.
DR PROSITE; PS00107; PROTEIN KINASE ATP; UNKNOWN_1.
DR PROSITE; PS00142; ZINC PROTEASE; UNKNOWN_1.
DR KW Hydrolyase; Metalloprotease; Protease; Zinc.
NON TER 1
SQ SEQUENCE 837 AA; 95505 MW; 3681FA41EA66FB3E CRC64;

Query Match 64.5%; Score 40; DB 2; Length 837;
Best Local Similarity 70.0%; Pred. No. 1.4e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 MPQNFYKLPQ 10
Db 255 MPREFYVLPQ 264

RESULT 10

Q6UEA5 ID Q6UEA5 PRELIMINARY; PRT; 902 AA.
AC Q6UEA5; 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Zinc metalloproteinase 6.
OS Ancylostoma ceylanicum.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;
OC Ancylostomatidae; Ancylostomatinae; Ancylostoma.
OX NCBI_TaxID=53326;
RN [1]
RP SEQUENCE FROM N.A.
RA Wisniewski M., Jaros S., Cappello M., Wedrychowicz H.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBSJ databases.

DR EMBL; AY371701; AA075756.1; --
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0008237; F:metalloproteinase activity; IEA.
DR GO; GO:0004245; F:neprilysin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR002052; N6 Mase.
DR InterPro; IPR000718; Peptidase_M13.
DR InterPro; IPR008753; Peptidase_M13_N.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR Pfam; PF01431; Peptidase_M13; 1.
DR Pfam; PF05649; Peptidase_M13_N; 1.
DR PRINTS; PR00786; NEPRILYSIN.
DR PROSITE; PS00092; N6 Mase; UNKNOWN_1.
DR PROSITE; PS00142; ZINC PROTEASE; UNKNOWN_1.
KW Hydrolyase; Metalloprotease; Protease; Zinc.
SQ SEQUENCE 902 AA; 101872 MW; 0C9294E77851D25D CRC64;

Query Match 64.5%; Score 40; DB 2; Length 902;
Best Local Similarity 70.0%; Pred. No. 1.5e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 MPQNFYKLPQ 10
Db 333 MPREFYVLPQ 342

RESULT 11

O67490 ID O67490 PRELIMINARY; PRT; 247 AA.
AC O67490; 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE ABC transporter.
GN Name=abcrl0; OrderedLocusNames=AQ_1531;
OS Aquifex aeolicus.
OC Bacteria; Aquificae; Aquificales; Aquificaceae; Aquifex.
OX NCBI_TaxID=63363;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VF5;
RX MEDLINE=98196666; PubMed=9537320; DOI=10.1038/32831;
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
Graham D.E., Overbeek R., Shead M.A., Keller M., Aujay M., Huber R.,
Feldman R.A., Short J.M., Olsen G.J., Swanson R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex
aeolicus".
RL Nature 392:353-358(1998).
CC -1- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AE000744; AAC07448.1; --
DR PIR; H70432; H70432.
DR HGSP; P02915; 1B0U.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:000524; F:ATP binding; IEA.
DR GO; GO:0042826; F:ATPase activity, coupled to transmembrane m. . .; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR InterPro; IPR006810; P:transport; IEA.
DR InterPro; IPR003593; AAA ATPase.
DR Pfam; PF00005; ABC tran; 1.
DR ProDom; PD000006; ABC transporter; 1.
DR SMART; SM00382; AAA; 1.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 1.
KW ATP-binding; Complete proteome.
SQ SEQUENCE 247 AA; 27535 MW; 8C843C4C2F2AFD5F CRC64;

Query Match 62.9%; Score 39; DB 2; Length 247;
Best Local Similarity 66.7%; Pred. No. 60;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 2 PQNFYKLPQ 10
Db 130 PEDFYKVPQ 138

```

RESULT 12
SYD_CHLMU
ID SYD_CHLMU STANDARD; PRT; 582 AA.
AC Q9PJUK;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Aspartyl-tRNA synthetase [EC 6.1.1.12] (Aspartate--tRNA ligase)
DE (AspRS).
GN Name=asps; OrderedLocusNames=TC0829;
OS Chlamydia muridarum.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83560;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Mopn / N199;
RA MEDLINE=20150255; PubMed=10684935; DOI=10.1093/nar/28.6.1397;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J.D., Utterback T.R., Berry K.J.,
RA Baas S., Linher K.D., Weidman J.F., Khouri H.M., Craven B., Bowman C.,
RA Dodson R.J., Gwin M.L., Nelson W.C., DeBoy R.T., Kolonay J.F.,
RA McClarty G., Salzberg S.L., Eisen J.A., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
RT pneumoniae AR39."
RL Nucleic Acids Res. 28:1397-1406(2000).
CC -1- CATALYTIC ACTIVITY: ATP + L-aspartate + tRNA(Asp) = AMP +
CC diphosphate + L-aspartyl-tRNA(Asp).
CC -1- SUBUNIT: Homodimer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: Belongs to the class-II aminoacyl-tRNA synthetase
CC family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; AE002349; AAF39629.1; -.
CC PIR; E81660; E81660.
CC HSP; P36419; I10W.
CC TIGR; TC0829; -.
CC HAMAP; MF_00044; -.
CC InterPro; IPR004524; Asps_bact.
CC InterPro; IPR004115; GAD.
CC InterPro; IPR008994; Nucleic acid OB.
CC InterPro; IPR004364; tRNA-synt_2.
CC InterPro; IPR002312; tRNA-synt_asp.
CC InterPro; IPR004365; tRNA anti.
CC Pfam; PF02938; GAD; 1.
CC Pfam; PF00152; tRNA-synt_2; 2.
CC Pfam; PF01336; tRNA anti_1.
CC PRINTS; PR01042; TRNASYNTHASP.
CC TIGRFAMs; TIGR00459; asps_bact; 1.
CC PROSITE; PS50862; AA TRNA LIGASE II; 1.
CC Aminoacyl-tRNA synthetase; ATP-binding; Complete proteome; Ligase;
CC Protein biosynthesis.
CC SEQUENCE 582 AA; 66386 MW; 492FC82BB710D420 CRC64;

Query Match 62.9%; Score 39; DB 1; Length 582;
Best Local Similarity 77.8%; Pred. No. 1.5e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 2 PQNFYKLPQ 10
Db 190 PQNFYKLPQ 198

RESULT 13
SYD_CHLTR
ID SYD_CHLTR STANDARD; PRT; 582 AA.
AC O84546;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Aspartyl-tRNA synthetase [EC 6.1.1.12] (Aspartate--tRNA ligase)
DE (AspRS).
GN Name=asps; OrderedLocusNames=CT542;
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=813;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D/UW-3/Cx;
RA MEDLINE=9900809; PubMed=9784136; DOI=10.1126/science.282.5389.754;
RA Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L.,
RA Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,
RA Davis R.W.;
RT "Genome sequence of an obligate intracellular pathogen of humans:
RT Chlamydia trachomatis."
RL Science 282:754-759(1998).
CC -1- CATALYTIC ACTIVITY: ATP + L-aspartate + tRNA(Asp) = AMP +
CC diphosphate + L-aspartyl-tRNA(Asp).
CC -1- SUBUNIT: Homodimer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: Belongs to the class-II aminoacyl-tRNA synthetase
CC family.
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; AE001325; AAC68144.1; -.
CC PIR; G71500; G71500.
CC HSP; P36419; I10W.
CC HAMAP; MF_00044; -.
CC InterPro; IPR004524; Asps_bact.
CC InterPro; IPR004115; GAD.
CC InterPro; IPR008994; Nucleic acid OB.
CC InterPro; IPR004364; tRNA-synt_2.
CC InterPro; IPR002312; tRNA-synt_asp.
CC InterPro; IPR004365; tRNA anti.
CC InterPro; IPR006195; tRNA_ligase_II.
CC Pfam; PF02938; GAD; 1.
CC Pfam; PF00152; tRNA-synt_2; 2.
CC Pfam; PF01336; tRNA anti_1.
CC PRINTS; PR01042; TRNASYNTHASP.
CC TIGRFAMs; TIGR00459; asps_bact; 1.
CC PROSITE; PS50862; AA TRNA LIGASE II; 1.
CC Aminoacyl-tRNA synthetase; ATP-binding; Complete proteome; Ligase;
CC Protein biosynthesis.
CC SEQUENCE 582 AA; 66227 MW; E8D2EBB2785EFABA CRC64;

Query Match 62.9%; Score 39; DB 1; Length 582;
Best Local Similarity 77.8%; Pred. No. 1.5e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 2 PQNFYKLPQ 10
Db 190 PQNFYKLPQ 198

RESULT 14
SYD_BUCAP
ID SYD_BUCAP STANDARD; PRT; 584 AA.
AC P81432;
DT 15-DEC-1998 (Rel. 37, Created)

```

DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Aspartyl-tRNA synthetase (EC 6.1.1.12) (Aspartate--tRNA ligase)
 DE (AspRS).
 GN Name=asps; OrderedLocusNames=BUG306;
 OS Buchnera aphidicola (subsp. Schizaphis graminum).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Buchnera.
 OX NCBI_TaxID=98794;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Zhao M.L., Baumann P.;
 RA "Nucleotide sequence of a DNA fragment from Buchnera aphidicola (Aphid
 RT endosymbiont) containing the genes asps-trxB-serS-serC-aroA-tpsa-hind-
 RT tpiA".
 RL Curr. Microbiol. 35:68-69(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22084549; PubMed=12089438; DOI=10.1126/science.1071278;
 RA Tamás I., Klasson L., Cambaeck B., Naeslund A.K., Eriksson A.-S.,
 RA Wernegreen J.J., Sandstrom J.P., Moran N.A., Andersson S.G.E.;
 RL "50 million years of genomic stasis in endosymbiotic bacteria".
 RL Science 296:2376-2379(2002).
 CC -1- CATALYTIC ACTIVITY: ATP + L-aspartate + tRNA(Asp) = AMP +
 CC diphosphate + L-aspartyl-tRNA(Asp).
 CC -1- SUBUNIT: Homodimer (By similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: Belongs to the class-II aminoacyl-tRNA synthetase
 CC family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; L43549; AAC05432.1; -.
 DR EMBL; AE014107; AAM67850.1; -.
 DR HSSP; P13030; 1BBW.
 DR HAWAP; MF 00044; -; 1.
 DR InterPro; IPR004524; Asps_bact.
 DR InterPro; IPR004115; GAD.
 DR InterPro; IPR008994; Nucleic_acid_OB.
 DR InterPro; IPR004364; tRNA-synt_2.
 DR InterPro; IPR002312; tRNA-synt_2.
 DR InterPro; IPR004365; tRNA-anti.
 DR InterPro; IPR006195; tRNA_ligase_II.
 DR Pfam; PF02938; GAD; 1.
 DR Pfam; PF00152; tRNA-synt_2; 2.
 DR Pfam; PF01336; tRNA-anti; 1.
 DR PRINTS; PR01042; TRNASYNTHASP.
 DR TIGRFAMS; TIGR00459; asps_bact; 1.
 DR PROSITE; PS00862; AA TRNA_LIGASE II; 1.
 DR Aminoacyl-tRNA synthetase; ATP-binding; Complete proteome; Ligase;
 KW Protein biosynthesis.
 FT DOMAIN 79 84
 FT SEQUENCE 584 AA; 67824 MW; F318DFA1654780E7 CRC64;
 SQ
 Query Match 62.9%; Score 39; DB 1; Length 584;
 Best Local Similarity 77.8%; Pred. No. 1.5e+02;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 2 PONFYKLPQ 10
 DB 182 PGNFYALPQ 190
 | | | | |
 RESULT 15
 Q6YR88 PRELIMINARY; PRT; 586 AA.
 ID Q6YR88
 AC Q6YR88;

DT 05-JUL-2004 (TRENBLrel. 27, Created)
 DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)
 DE Aspartyl-tRNA synthetase.
 DE Aspartyl-tRNA synthetase.
 GN Name=asps; OrderedLocusNames=PAM128;
 OS Onion yellow phytoplasma.
 OC Bacteria; Firmicutes; Mollicutes; Achleoplasmatales;
 OC Achleoplasmataceae; Candidatus Phytoplasma.
 OX NCBI_TaxID=100379;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN=OY-N;
 RX PubMed=14661021; DOI=10.1038/ng1277;
 RA Oshima K., Kikizawa S., Nishigawa H., Jung H.-Y., Wei W., Suzuki S.,
 RA Arashida R., Nakata D., Miyata S.-I., Ugaki M., Namba S.;
 RL "Reductive evolution suggested from the complete genome sequence of a
 RL plant-pathogenic phytoplasma".
 RL Nat. Genet. 36:27-29(2004).
 CC -1- SIMILARITY: Belongs to class-II aminoacyl-tRNA synthetase family.
 CC EMBL; AF006628; BAD04213.1; -.
 DR HSSP; P13030; 1BBW.
 DR GO; GO:0005737; Cytoplasm; IEA.
 DR GO; GO:0004815; F:aspartate-tRNA ligase activity; IEA.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0003676; F:nucleic acid binding; IEA.
 DR GO; GO:0004222; P:aspartyl-tRNA aminoacylation; IEA.
 DR InterPro; IPR004524; Asps_bact.
 DR InterPro; IPR008994; Nucleic_acid_OB.
 DR InterPro; IPR004364; tRNA-synt_2.
 DR InterPro; IPR002312; tRNA-synt_2.
 DR InterPro; IPR006195; tRNA_ligase_II.
 DR Pfam; PF00152; tRNA-synt_2; 1.
 DR PRINTS; PR01042; TRNASYNTHASP.
 DR TIGRFAMS; TIGR00459; asps_bact; 1.
 DR PROSITE; PS00862; AA TRNA_LIGASE II; 1.
 DR ATP-binding; Aminoacyl-tRNA synthetase; Complete proteome; Ligase;
 KW Protein biosynthesis.
 SQ SEQUENCE 586 AA; 67520 MW; 7CC2D6920BA74B49 CRC64;
 Query Match 62.9%; Score 39; DB 2; Length 586;
 Best Local Similarity 77.8%; Pred. No. 1.5e+02;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 2 PONFYKLPQ 10
 DB 186 PGNFYALPQ 194
 | | | | |
 Search completed: April 12, 2005, 08:05:46
 Job time : 54.2 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 12, 2005, 07:49:31 ; Search time 84 Seconds
(without alignments)
69.064 Million cell updates/sec

Title: US-10-691-330-7

Perfect score: 81

Sequence: 1 VLEMKPPPPQETVT 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq_16Dec04:*

1: geneseqp1980s:*

2: geneseqp1990s:*

3: geneseqp2000s:*

4: geneseqp2001s:*

5: geneseqp2002s:*

6: geneseqp2003as:*

7: geneseqp2003bs:*

8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	81	100.0	15	4	AAB72506 Colostrin
2	81	100.0	15	4	AAB59312 Ewe colos
3	81	100.0	15	4	AAB72252 Colostrin
4	81	100.0	15	4	AAB72538 Colostrin
5	81	100.0	15	5	AAB14583 Neural ce
6	81	100.0	15	5	AAM51042 Colostrin
7	81	100.0	15	5	AAB20234 Colostrin
8	81	100.0	15	8	ADN60301 Constitue
9	81	100.0	15	8	ADN74388 Ovine col
10	81	100.0	16	4	AAB59343 Ewe colos
11	51	63.0	164	3	AAG38918 Arabidops
12	51	63.0	218	3	AAG19665 Arabidops
13	51	63.0	239	3	AAG19664 Arabidops
14	51	63.0	268	3	AAG38917 Arabidops
15	51	63.0	289	3	AAG38916 Arabidops
16	51	63.0	289	4	AAB02472 Arabidops
17	51	63.0	289	6	ADB23136 Environme
18	51	63.0	289	7	ADB37165 Plant.yie.
19	51	63.0	289	8	ADI43551 Plant tra
20	51	63.0	289	8	ADO01901 Thalecres
21	51	63.0	289	8	ADN73653 Thale cre
22	51	63.0	319	7	ABO43111 A. thalia
23	51	63.0	319	7	ADC46623 Thalecres
24	51	63.0	319	7	ADD55790 Thalecres
25	49	60.5	180	8	ADO62175 Transcrip

ALIGNMENTS

RESULT 1

AAB72506
ID AAB72506 standard; peptide; 15 AA.

XX AAB72506;

XX AC (first entry)

DT 09-MAY-2001 (first entry)

DE Colostrin peptide #7.

XX Dermatology; oxidative stress regulator; colostrinin.

XX Unidentified.

XX W0200112650-A2.

XX 22-FEB-2001.

XX 17-AUG-2000; 2000WO-US022665.

XX 17-AUG-1999; 99US-0149310P.

XX (TEXA) UNIV TEXAS SYSTEM.

XX Stanton GJ, Hughes TK, Boldogh I;

DR WPI; 2001-218342/22.

XX Modulating oxidative stress level in a cell, involves contacting the cell with an oxidative stress regulator selected from colostrinin, its constituent peptide, analog or their combinations.

XX Claim 6; Page 25; 48pp; English.

XX The present invention relates to a method for modulating the oxidative stress level in a cell or a patient, comprising contacting the cell with, or administering to the patient, an oxidative stress regulator selected from colostrinin, or its constituent peptide (e.g. the present peptide), to change the level of an oxidising species in the cell. The method can be used to treat oxidative damage to skin, by decreasing or preventing an increase in the level of damage to a biomolecule of the patient

XX Sequence 15 AA;

Query Match 100.0%; Score 81; DB 4; Length 15;

Best Local Similarity 100.0%; Pred. No. 9.7e-05;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Aao06516 Human pol
Abg19150 Novel hum
Abo80666 Pseudomon
Adn20996 Bacterial
Aam17645 Peptide #
Abb36665 Peptide #
Aam30163 Peptide #
Abb31452 Peptide #
Abb22001 Protein #
Aam69824 Human bon
Aam57428 Human bra
Aag51515 Human liv
Aam05303 Peptide #
Abg39451 Human pep
Abb67485 Drosophil
Abg08125 Novel hum
Aag08618 Arabidops
Add30054 Plant yie
Ade31443 Plant yie
Adi44347 Plant tra

26 47 58.0 218 4 AAO06516
27 46 56.8 141 4 ABG19150
28 45 56.8 300 7 ABO80666
29 46 56.8 1196 8 ADN20996
30 45 55.6 78 4 AAM17645
31 45 55.6 78 4 ABB36665
32 45 55.6 78 4 AAM30163
33 45 55.6 78 4 ABB31452
34 45 55.6 78 4 ABB22001
35 45 55.6 78 4 AAM69824
36 45 55.6 78 4 AAM57428
37 45 55.6 78 4 AAG51515
38 45 55.6 78 4 AAM05303
39 45 55.6 78 5 ABG39451
40 45 55.6 114 4 ABB67485
41 45 55.6 366 4 ABB08125
42 44 54.3 166 3 AAG08618
43 44 54.3 166 7 ADD30054
44 44 54.3 166 7 ADE31443
45 44 54.3 166 8 ADI44347

Tue Apr 12 09:04:59 2005

dementia; neurodegenerative disease; Alzheimer's disease; psychosis; neurosis; infection.

Synthetic.

WO200111937-A2.

22-FEB-2001.

17-AUG-2000; 2000WO-US022818.

17-AUG-1999; 99US-0149311P.

(TEXA) UNIV TEXAS SYSTEM.

(REGE-) REGEN THERAPEUTICS PLC.

Stanton GJ, Hughes TK, Boldogh I, Georgiades J;

WPI; 2001-202804/20.

Inducing a cytokine and modulating an immune response, useful for treating central nervous system diseases and bacterial and viral infections, comprises administering colostrinin as an immunological regulator.

Claim 1; Page 34; 50pp; English.

Sequences AAB72246 - AAB72275 represent peptides derived from colostrinin, a proline rich polypeptide aggregate contained in colostrum. The peptides have immune response modulatory activity, and are capable of inducing cytokines. Colostrinin and its derived peptides are useful for inducing cytokine production, for modulating an immunological response and for inducing blood cell proliferation. The peptides are useful in the treatment of disorders of the central nervous system, neurological disorders, mental disorders, dementia, neurodegenerative diseases, Alzheimer's disease, motor neurons disease, psychosis, neurosis, chronic disorders of the immune system, bacterial and viral infections and acquired immunological deficiencies

Sequence 15 AA;

Query Match 100.0%; Score 81; DB 4; Length 15; Best Local Similarity 100.0%; Pred. No. 9.7e-05; Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 VLEMKFPPPPQETVT 15
Dy 1 VLEMKFPPPPQETVT 15

RESULT 4
AAB72538

ID AAB72538 standard; peptide; 15 AA.

XX AAB72538;

AC AAB72538;

DT 09-MAY-2001 (first entry)

XX Colostrinin peptide #7.

XX Neuroprotective; neural cell differentiation regulator; colostrinin;

XX Colostrum.

XX Unidentified.

XX WO200112651-A2.

XX 22-FEB-2001.

XX 17-AUG-2000; 2000WO-US022774.

XX 17-AUG-1999; 99US-0149633P.

XX

OS OS

PN PN

PD PD

PF PF

PR PR

XX XX

XX XX

PA PA

PA PA

PI PI

XX XX

DR DR

XX XX

PT PT

PT PT

PT PT

XX XX

PS PS

XX XX

CC CC

CC CC

CC CC

CC CC

CC CC

CC CC

CC CC

CC CC

CC CC

CC CC

CC CC

CC CC

CC CC

CC CC

CC CC

CC CC

CC CC

CC CC

CC CC

CC CC

CC CC

OS OS

PN PN

PD PD

PF PF

PR PR

XX XX

XX XX

PA PA

PA PA

PI PI

XX XX

DR DR

XX XX

PT PT

PT PT

PT PT

XX XX

PS PS

XX XX

CC CC

CC CC

CC CC

CC CC

CC CC

CC CC

CC CC

CC CC

CC CC

CC CC

CC CC

CC CC

CC CC

CC CC

CC CC

CC CC

CC CC

CC CC

CC CC

CC CC

CC CC

Oy 1 VLEMKFPPPPQETVT 15
Dy 1 VLEMKFPPPPQETVT 15

RESULT 2

AAB59312

ID AAB59312 standard; peptide; 15 AA.

XX AAB59312;

AC AAB59312;

DT 21-MAR-2001 (first entry)

XX Ewe colostrinin peptide fragment A-3.

XX Sheep; colostrinin; proline rich polypeptide; colostrum; immune disorder;

XX central nervous system disorder; dietary supplement; beta-amyloid plaque.

XX Ovis sp.

XX WO200075173-A2.

XX 14-DEC-2000.

XX 02-JUN-2000; 2000WO-GB002128.

XX 02-JUN-1999; 99GB-00012852.

XX (REGE-) REGEN THERAPEUTICS PLC.

XX Georgiades JA;

XX WPI; 2001-071058/08.

XX Peptides having an N-terminal amino acid sequence isolated from

XX colostrinin for treating e.g. disorders of the central nervous system and

XX immune system, viral and bacterial infections, and diseases characterized

XX by amyloid plaques.

XX Claim 7; Page 27; 63pp; English.

XX The present invention provides the sequences of a number of peptides

XX found in ewe's colostrinin. Colostrinin is the proline-rich polypeptide

XX fragment of colostrum. These peptides can be used in the treatment of

XX central nervous system disorders such as senile dementia, Parkinson's

XX disease, Alzheimer's disease, psychosis and neurosis, immune system

XX disorders such as bacterial and viral infections, to improve the

XX development of a child's immune system, as a dietary supplement, and to

XX promote the dissolution of beta-amyloid plaques

XX Sequence 15 AA;

Query Match 100.0%; Score 81; DB 4; Length 15;

Best Local Similarity 100.0%; Pred. No. 9.7e-05;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 VLEMKFPPPPQETVT 15

Dy 1 VLEMKFPPPPQETVT 15

RESULT 3

AAB72252

ID AAB72252 standard; peptide; 15 AA.

XX AAB72252;

XX 14-MAY-2001 (first entry)

XX Colostrinin derived cytokine inducing peptide SEQ ID 7.

XX Colostrinin; immune response; cytokine; blood cell proliferation;

XX central nervous system disorder; neurological disorder; mental disorder;

XX

PA (TEXA) UNIV TEXAS SYSTEM.
XX
PI Boldog I;
XX
XX WPI; 2001-226545/23.
XX
XX Use of colostrin, its constituent peptide or analog as a neural cell
PT regulator, for promoting neural cell differentiation and treating damaged
PT neural cells in a patient.
XX
XX Claim 6; Page 21; 35pp; English.
XX
XX The present invention relates to a method for promoting neural cell
CC differentiation and treating damaged neural cells, using colostrin and
CC colostrin constituent peptides (e.g. the present peptide) as a neural
CC cell regulator. Colostrin is a polypeptide complex found in colostrum
XX
XX
SQ Sequence 15 AA;

Query Match 100.0%; Score 81; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 9.7e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLEMKFPPPPQETVT 15
| | | | | | | | | | | | | | |
Db 1 VLEMKFPPPPQETVT 15

RESULT 5
AAO14583
ID AAO14583 standard; peptide; 15 AA.
XX
AC AAO14583;
XX
XX 27-MAY-2002 (first entry)
DT
DE Neural cell regulatory colostrin peptide 7.
XX
XX Neural cell differentiation; neural cell regulator; colostrin peptide;
KW neural cell formation; proline-rich polypeptide aggregate; colostrum;
KW neural cell treatment.
XX
XX Unidentified.
XX
XX Key Location/Qualifiers
FH Modified-site 15
FT /note= "Optional C-terminal amide"
XX
XX WO200213851-A1.
XX
XX 21-FEB-2002.
XX
XX 17-AUG-2000; 2000WO-US022777.
XX
XX 17-AUG-2000; 2000WO-US022777.
XX
XX (TEXA) UNIV TEXAS SYSTEM.
XX
XX Boldog I, Stanton JG, Hughes TK;
XX
XX WPI; 2002-269152/31.
XX
XX Promoting cell differentiation in a patient involves use of blood cell
PT regulator selected from colostrin, its constituent peptide and/or
PT analog.
XX
XX Claim 7; Page 21; 37pp; English.
XX
XX The invention comprises a method for promoting cell differentiation (e.g.
CC neural cell differentiation). The method involves contacting cells with a
CC neural cell regulator (i.e. a colostrin peptide) in order to change the
CC cells in morphology to form neural cells. Colostrin is a proline-rich
CC polypeptide aggregate that is present in colostrum. The method of the

CC invention is useful for promoting the differentiation of cells and for
CC treating damaged neural cells in a patient. The present amino acid
CC sequence represents a specifically claimed colostrin peptide used in
CC the method of the invention
XX
XX Sequence 15 AA;

Query Match 100.0%; Score 81; DB 5; Length 15;
Best Local Similarity 100.0%; Pred. No. 9.7e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLEMKFPPPPQETVT 15
| | | | | | | | | | | | | | |
Db 1 VLEMKFPPPPQETVT 15

RESULT 6
AAM51042
ID AAM51042 standard; peptide; 15 AA.
XX
AC AAM51042;
XX
XX 30-MAY-2002 (first entry)
DT
DE Colostrin constituent peptide.
XX
XX Colostrin; colostrum; immunomodulator; cardiovascular;
KW blood cell regulator; cytokine inducer; human.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FH Modified-site 15
FT /note= "Optional C-terminal amidation"
XX
XX WO200213849-A1.
XX
XX 21-FEB-2002.
XX
XX 17-AUG-2000; 2000WO-US022775.
XX
XX 17-AUG-2000; 2000WO-US022775.
XX
XX (TEXA) UNIV TEXAS SYSTEM.
XX (REG-) REGEN THERAPEUTICS PLC.
XX
XX Stanton GJ, Hughes TK, Boldog I, Georgiades J;
PI
XX WPI; 2002-269150/31.
XX
XX Modulation of blood cell proliferation in a patient involves use of blood
PT cell regulator selected from colostrin, its constituent peptide and/or
PT analog.
XX
XX Claim 1; Page 34; 54pp; English.
XX
XX The present sequence is that of a colostrin constituent peptide that is
CC preferred for use as an immunological regulator and as a blood cell
CC regulator in claimed methods of the invention. Methods are claimed for:
CC inducing a cytokine in a cell by contact with an immunological regulator,
CC where the cell is present in a cell culture, a tissue, an organ or an
CC organism, and the cell is mammalian, including human; modulating an
CC immune response in a cell by contact with the immunological regulator
CC under conditions effective to induce a cytokine; modulating an immune
CC response in a patient by administering an immunological regulator under
CC conditions effective to induce a cytokine, where the immunological
CC regulator is administered topically or as part of a dietary supplement,
CC and where the immune response is specific or non specific, an interferon
CC response or an antibody response; modulating blood cell proliferation by
CC contacting blood cells with a blood cell regulator, where the blood cells
CC are present in a cell culture or an organism, are mammalian or human, and
CC where the blood cells are increased in number or differentiated; and a
CC method for modulating blood cell proliferation in a patent. A claimed

CC cytokine-inducing composition comprises a pharmaceutical carrier and an
CC active agent such as the present peptide. Cytokines induced by this
CC peptide in human leucocyte cultures include interferon-gamma, tumour
CC necrosis factor-alpha, interleukin-4, interleukin-6 and interleukin-10
XX
SQ Sequence 15 AA;
Query Match 100.0%; Score 81; DB 5; Length 15;
Best Local Similarity 100.0%; Pred. No. 9.7e-05; Indels 0; Gaps 0;
Matches 15; Conservative 0; Mismatches 0;
QY 1 VLEMKFPPPPQETVT 15
Db 1 VLEMKFPPPPQETVT 15
RESULT 7
AAE20234
ID AAE20234 standard; peptide; 15 AA.
XX
AC AAE20234;
XX
DT 18-JUN-2002 (first entry)
XX
DE Colostatin constituent peptide #7.
KW Blood cell regulator; colostatin; constituent peptide; oxidative stress;
KW therapy; oxidative damage; skin; aging; wound healing; cell replacement;
KW tissue; organ; cosmetic procedure; repair; regeneration; preservation;
KW transplantation; implantation; dermatological; vulnary.
XX
OS Unidentified.
XX
FH Key Location/Qualifiers
FT Modified-site 15 /note= "Optionally C-terminal amide"
FT
PN WO200213850-A1.
XX
PD 21-FEB-2002.
XX
PF 17-AUG-2000; 2000WO-US022776.
XX
PR 17-AUG-2000; 2000WO-US022776.
XX
PA (TEXA) UNIV TEXAS SYSTEM.
XX
PI Stanton GJ, Hughes TK, Boldogh I;
XX
DR WPI; 2002-269151/31.
XX
PT Composition useful for the modulation of blood cell proliferation in a
PT patient comprises a blood cell regulator selected from colostatin, its
PT constituent peptide and/or analog.
XX
PS Claim 6; Page 25; 51pp; English.
XX
CC The invention relates to a composition which comprises a blood cell
CC regulator selected from colostatin, its constituent peptide and/or
CC analogue. The invention is used for modulating the oxidative stress level
CC in a cell e.g. mammalian or human cell present in a cell culture, tissue,
CC organ, or organism; or for treating oxidative damage to the skin of a
CC patient e.g. animal or human; to modulate oxidative stress during/ after
CC a premature birth or normal birth, preventing/delaying aging in a
CC patient, enhancing wound healing, and the reduction of side effects of
CC cosmetic procedures. The method changes the level of an oxidising species
CC in the cell, such as decreases or prevents increase in the level of
CC damage to a biomolecule of the patient selected from DNA, protein and/or
CC lipid, compared to the same conditions when the oxidative stress
CC regulator is not present. The modulation of oxidative stress results in
CC enhanced repair, regeneration, and replacement of cells, tissues and
CC organs (e.g. kidney, liver, pancreas, skin, and the other internal and
CC external organs), as well as enhanced preservation of such organs for

CC transplantation, implantation, or scientific research. The present
CC sequence is a colostatin constituent peptide
XX
SQ Sequence 15 AA;
Query Match 100.0%; Score 81; DB 5; Length 15;
Best Local Similarity 100.0%; Pred. No. 9.7e-05; Indels 0; Gaps 0;
Matches 15; Conservative 0; Mismatches 0;
QY 1 VLEMKFPPPPQETVT 15
Db 1 VLEMKFPPPPQETVT 15
RESULT 8
ADN60301
ID ADN60301 standard; peptide; 15 AA.
XX
AC ADN60301;
XX
DT 29-JUL-2004 (first entry)
XX
DE Constituent peptide of colostatin SEQ ID NO:7.
KW modulator; colostatin; intracellular signaling molecule modulator;
KW 4HNE-mediated lipid peroxidation down regulation; apoptosis inhibition;
KW DNA damage; beta-amyloid; retinoic acid; cytostatic; 4HNE inhibitor;
KW 4HNE-protein adduct formation reduction;
KW 4HNE-mediated glutathione depletion inhibition; p53 inhibition;
KW c-Jun NH2-terminal kinase inhibition.
XX
OS Synthetic.
XX
PN WO2004037851-A2.
XX
PD 06-MAY-2004.
XX
PF 22-OCT-2003; 2003WO-US033423.
XX
PR 22-OCT-2002; 2002US-0420369P.
XX
PA (TEXA) UNIV TEXAS SYSTEM.
PA (BOLD/) BOLDOGH I.
PA (STAN/) STANTON J G.
PA (GEOR/) GEORGIADIS J A.
PA (HUGH/) HUGHES T K.
PA (KRUI/) KRUIZEL M.
XX
PI Boldogh I, Stanton JG, Georgiades JA, Hughes TK, Kruzel M;
XX
DR WPI; 2004-365494/34.
XX
PT Use of colostatin for e.g. modulating an intracellular signaling
PT molecule in a cell, down regulating 4HNE-mediated lipid peroxidation in a
PT cell, inhibiting apoptosis in a cell, or protecting against DNA damage in
PT a cell.
XX
PS Claim 6; SEQ ID NO 7; 46pp; English.
XX
CC The present invention describes the use of a modulator selected from
CC colostatin, its constituent peptide, its active analogue, and a
CC combination of these, for modulating an intracellular signaling molecule
CC in a cell, down regulating 4HNE-mediated lipid peroxidation in a cell,
CC inhibiting apoptosis in a cell, protecting against DNA damage in a cell,
CC and reducing the toxic effect of beta-amyloid or retinoic acid on a cell.
CC The modulator has cytostatic activity, and can be used as a 4HNE
CC inhibitor. The modulator is useful in the manufacture of a medicament for
CC reducing 4HNE-protein adduct formation, inhibiting 4HNE-mediated
CC glutathione depletion, inhibiting 4HNE-induced activation of p53 protein,
CC and/or inhibiting 4HNE-induced activation of c-Jun NH2-terminal kinases.
CC Colostatin, or its constituent peptide or active analogue is useful for
CC inhibiting apoptosis in a cell, protecting against DNA damage in a cell,
CC or reducing the toxic effect of beta-amyloid or retinoic acid on a cell.

CC The present sequence represents a synthetic constituent peptide of
 CC colostrinin, which can be used as a modulator in the present invention.
 XX
 SQ Sequence 15 AA;

Query Match 100.0%; Score 81; DB 8; Length 15;
 Best Local Similarity 100.0%; Pred. No. 9.7e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLEMKFPPPPQETVT 15
 |||||
 Db 1 VLEMKFPPPPQETVT 15

RESULT 9
 ADS74388
 ID ADS74388 standard; peptide; 15 AA.

XX AC ADS74388;

XX DT 16-DEC-2004 (first entry)

XX DE Ovine colostrinin peptide.

XX KW Colostrum; colostrinin; sheep; peptide purification.

XX OS Ovis aries.

XX FN WO2004081038-A1.

XX PD 23-SEP-2004.

XX PF 10-MAR-2004; 2004WO-GB001014.

XX PR 11-MAR-2003; 2003GB-00005552.

XX PR 08-MAR-2004; 2004GB-00005190.

XX PA (REGG-) REGEN THERAPEUTICS PLC.

XX PI Georgiades JA, Polanowski A, Wilusz T, Kruzel ML;

XX DR WPI; 2004-677519/66.

PT Recovering peptides such as colostrinin from mammalian colostrum, by
 PT mixing colostrum with alcohol to form alcohol phase containing peptides
 PT and precipitate, separating alcohol phase from precipitate, and
 PT recovering alcohol phase.

PS Example; SEQ ID NO 3; 41pp; English.

XX The present sequence is that of a peptide that can be recovered from
 CC ovine colostrinin using the method of the invention. The invention
 CC provides a method for the recovery of peptides (especially colostrinin)
 CC from colostrum in substantially pure, biologically active form and in
 CC high yield. The method involves mixing the colostrum with an alcohol to
 CC form an alcohol phase containing the colostrinin and a precipitate
 CC containing higher molecular weight caseins and other proteins. Best
 CC results are obtained using methanol or ethanol of at least 80%, and
 CC preferably up to 100%, purity. The alcohol phase is then separated from
 CC the precipitate, and the colostrinin is separated from the alcohol,
 CC preferably by evaporation, to form a colostrinin-rich phase, which is
 CC recovered. A precipitation agent, such as ammonium sulfate, may be added
 CC either to the alcohol phase or, preferably, to the colostrinin-rich phase
 CC to induce precipitation of the colostrinin peptides. The method is
 CC generally applicable to the separation of peptides from fluids containing
 CC higher molecular weight proteins, lipids, carbohydrates and/or nucleic
 CC acids. In an example from the invention, the antigenic profile of
 CC peptides recovered from sheep colostrum using the alcohol precipitation
 CC methods was determined by ELISA using antibodies prepared against 9
 CC synthetic peptides, including a peptide having the present sequence
 CC (denoted antigen class A-3).

XX Sequence 15 AA;

Query Match 100.0%; Score 81; DB 8; Length 15;
 Best Local Similarity 100.0%; Pred. No. 9.7e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLEMKFPPPPQETVT 15
 |||||
 Db 1 VLEMKFPPPPQETVT 15

RESULT 10
 AAB59343
 ID AAB59343 standard; peptide; 16 AA.

XX AC AAB59343;

XX DT 21-MAR-2001 (first entry)

XX DE Ewe colostrinin peptide fragment derived sequence #3.

XX KW Sheep; colostrinin; proline rich polypeptide; colostrum; immune disorder;
 KW central nervous system disorder; dietary supplement; beta-amyloid plaque.

XX OS Ovis sp.

XX FN WO200075173-A2.

XX PD 14-DEC-2000.

XX PF 02-JUN-2000; 2000WO-GB002128.

XX PR 02-JUN-1999; 99GB-00012852.

XX PA (REGG-) REGEN THERAPEUTICS PLC.

XX PI Georgiades JA;

XX DR WPI; 2001-071058/08.

XX PT Peptides having an N-terminal amino acid sequence isolated from
 PT colostrinin for treating e.g. disorders of the central nervous system and
 PT immune system, viral and bacterial infections, and diseases characterized
 PT by amyloid plaques.

XX PS Claim 8; Page 27; 63pp; English.

XX The present invention provides the sequences of a number of peptides
 CC found in ewe's colostrinin. Colostrinin is the proline-rich polypeptide
 CC fragment of colostrum. These peptides can be used in the treatment of
 CC central nervous system disorders such as senile dementia, Parkinson's
 CC disease, Alzheimer's disease, psychosis and neurosis, immune system
 CC disorders such as bacterial and viral infections, to improve the
 CC development of a child's immune system, as a dietary supplement, and to
 CC promote the dissolution of beta-amyloid plaques

XX SQ Sequence 16 AA;

Query Match 100.0%; Score 81; DB 4; Length 16;
 Best Local Similarity 100.0%; Pred. No. 0.0001;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLEMKFPPPPQETVT 15
 |||||
 Db 2 VLEMKFPPPPQETVT 16

RESULT 11
 AAG38918
 ID AAG38918 standard; protein; 164 AA.

XX AC AAG38918;

XX DT 18-OCT-2000 (first entry)

Tue Apr 12 09:04:59 2005

us-10-691-330-7.rag

```
XX Arabidopsis thaliana protein fragment SEQ ID NO: 48081.
DE
XX Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
KW
XX Arabidopsis thaliana.
XX
XX EPI033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-00301439.
XX
XX 25-FEB-1999; 99US-0121825P.
XX 05-MAR-1999; 99US-01231180P.
XX 09-MAR-1999; 99US-0123548P.
XX 23-MAR-1999; 99US-0125788P.
XX 25-MAR-1999; 99US-0126264P.
XX 29-MAR-1999; 99US-0126785P.
XX 01-APR-1999; 99US-0127462P.
XX 06-APR-1999; 99US-0128234P.
XX 08-APR-1999; 99US-0128714P.
XX 16-APR-1999; 99US-0129845P.
XX 19-APR-1999; 99US-0130077P.
XX 21-APR-1999; 99US-0130449P.
XX 23-APR-1999; 99US-0130510P.
XX 28-APR-1999; 99US-0130891P.
XX 30-APR-1999; 99US-0131449P.
XX 30-APR-1999; 99US-0132048P.
XX 30-APR-1999; 99US-0132407P.
XX 04-MAY-1999; 99US-0132484P.
XX 05-MAY-1999; 99US-0132485P.
XX 06-MAY-1999; 99US-0132486P.
XX 06-MAY-1999; 99US-0132487P.
XX 07-MAY-1999; 99US-0132863P.
XX 11-MAY-1999; 99US-0134256P.
XX 14-MAY-1999; 99US-0134218P.
XX 14-MAY-1999; 99US-0134219P.
XX 14-MAY-1999; 99US-0134221P.
XX 14-MAY-1999; 99US-0134370P.
XX 18-MAY-1999; 99US-0134768P.
XX 19-MAY-1999; 99US-0134941P.
XX 20-MAY-1999; 99US-0135124P.
XX 21-MAY-1999; 99US-0135353P.
XX 24-MAY-1999; 99US-0135623P.
XX 25-MAY-1999; 99US-0136021P.
XX 27-MAY-1999; 99US-0136392P.
XX 28-MAY-1999; 99US-0136782P.
XX 01-JUN-1999; 99US-0137222P.
XX 03-JUN-1999; 99US-0137528P.
XX 04-JUN-1999; 99US-0137502P.
XX 07-JUN-1999; 99US-0137724P.
XX 08-JUN-1999; 99US-0138094P.
XX 10-JUN-1999; 99US-0138540P.
XX 10-JUN-1999; 99US-0138847P.
XX 14-JUN-1999; 99US-0139119P.
XX 16-JUN-1999; 99US-0139453P.
XX 17-JUN-1999; 99US-0139492P.
XX 18-JUN-1999; 99US-0139454P.
XX 18-JUN-1999; 99US-0139455P.
XX 18-JUN-1999; 99US-0139456P.
XX 18-JUN-1999; 99US-0139457P.
XX 18-JUN-1999; 99US-0139458P.
XX 18-JUN-1999; 99US-0139459P.
XX 18-JUN-1999; 99US-0139460P.
XX 18-JUN-1999; 99US-0139461P.
XX 18-JUN-1999; 99US-0139462P.
XX 18-JUN-1999; 99US-0139463P.
XX 18-JUN-1999; 99US-0139750P.
XX 18-JUN-1999; 99US-0139763P.
XX 21-JUN-1999; 99US-0139817P.
XX 22-JUN-1999; 99US-0139899P.
XX 23-JUN-1999; 99US-0140353P.
XX 23-JUN-1999; 99US-0140354P.
XX 24-JUN-1999; 99US-0140695P.
XX 28-JUN-1999; 99US-0140823P.
XX 29-JUN-1999; 99US-0140991P.
XX 30-JUN-1999; 99US-0141287P.
XX 01-JUL-1999; 99US-0141842P.
XX 01-JUL-1999; 99US-0142154P.
XX 02-JUL-1999; 99US-0142055P.
XX 06-JUL-1999; 99US-0142390P.
XX 08-JUL-1999; 99US-0142803P.
XX 09-JUL-1999; 99US-0142920P.
XX 12-JUL-1999; 99US-0142977P.
XX 13-JUL-1999; 99US-0143542P.
XX 14-JUL-1999; 99US-0143624P.
XX 15-JUL-1999; 99US-0144005P.
XX 16-JUL-1999; 99US-0144085P.
XX 16-JUL-1999; 99US-0144086P.
XX 19-JUL-1999; 99US-0144325P.
XX 19-JUL-1999; 99US-0144331P.
XX 19-JUL-1999; 99US-0144332P.
XX 19-JUL-1999; 99US-0144333P.
XX 19-JUL-1999; 99US-0144334P.
XX 20-JUL-1999; 99US-0144335P.
XX 20-JUL-1999; 99US-0144352P.
XX 20-JUL-1999; 99US-0144632P.
XX 20-JUL-1999; 99US-0144884P.
XX 21-JUL-1999; 99US-0144814P.
XX 21-JUL-1999; 99US-0145086P.
XX 21-JUL-1999; 99US-0145088P.
XX 22-JUL-1999; 99US-0145085P.
XX 22-JUL-1999; 99US-0145087P.
XX 22-JUL-1999; 99US-0145089P.
XX 22-JUL-1999; 99US-0145192P.
XX 23-JUL-1999; 99US-0145145P.
XX 23-JUL-1999; 99US-0145218P.
XX 23-JUL-1999; 99US-0145224P.
XX 26-JUL-1999; 99US-0145276P.
XX 27-JUL-1999; 99US-0145913P.
XX 27-JUL-1999; 99US-0145918P.
XX 27-JUL-1999; 99US-0145919P.
XX 28-JUL-1999; 99US-0145951P.
XX 02-AUG-1999; 99US-0146386P.
XX 02-AUG-1999; 99US-0146388P.
XX 03-AUG-1999; 99US-0146389P.
XX 03-AUG-1999; 99US-0147038P.
XX 04-AUG-1999; 99US-0147204P.
XX 04-AUG-1999; 99US-0147302P.
XX 05-AUG-1999; 99US-0147192P.
XX 05-AUG-1999; 99US-0147260P.
XX 06-AUG-1999; 99US-0147303P.
XX 06-AUG-1999; 99US-0147416P.
XX 09-AUG-1999; 99US-0147493P.
XX 09-AUG-1999; 99US-0147935P.
XX 10-AUG-1999; 99US-0148171P.
XX 11-AUG-1999; 99US-0148319P.
XX 12-AUG-1999; 99US-0148341P.
XX 13-AUG-1999; 99US-0148565P.
XX 13-AUG-1999; 99US-0148684P.
XX 16-AUG-1999; 99US-0149368P.
XX 17-AUG-1999; 99US-0149175P.
XX 18-AUG-1999; 99US-0149426P.
XX 20-AUG-1999; 99US-0149722P.
XX 20-AUG-1999; 99US-0149723P.
XX 20-AUG-1999; 99US-0149929P.
XX 23-AUG-1999; 99US-0149902P.
XX 23-AUG-1999; 99US-0149930P.
XX 25-AUG-1999; 99US-0150566P.
XX 26-AUG-1999; 99US-0150884P.
XX 27-AUG-1999; 99US-0151065P.
XX 27-AUG-1999; 99US-0151066P.
```

```
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159239P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 26-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 63.0%; Score 51; DB 3; Length 164;
Best Local Similarity 69.2%; Pred. No. 17;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VLEMKFPFPQET 13
Db 59 VTENVMPFPQOT 71

RESULT 12
AAG19665
ID AAG19665 standard; protein; 218 AA.
XX AC AAG19665;
XX DT
XX DE 17-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 21552.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX KW termination sequence.
XX
```

```
OS Arabidopsis thaliana.
XX XX
PN EP1033405-A2.
XX XX
PD 06-SEP-2000.
XX XX
XX 25-FEB-2000; 2000EP-00301439.
XX XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 23-MAR-1999; 99US-0123548P.
PR 25-MAR-1999; 99US-0125788P.
PR 29-MAR-1999; 99US-0126264P.
PR 01-APR-1999; 99US-0126785P.
PR 06-APR-1999; 99US-0127462P.
PR 08-APR-1999; 99US-0128234P.
PR 16-APR-1999; 99US-0128714P.
PR 19-APR-1999; 99US-0129845P.
PR 21-APR-1999; 99US-0130077P.
PR 23-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132407P.
PR 05-MAY-1999; 99US-0132484P.
PR 06-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
```

PR	30-JUN-1999;	99US-0141287P.
PR	01-JUL-1999;	99US-0141842P.
PR	01-JUL-1999;	99US-0142154P.
PR	02-JUL-1999;	99US-0142055P.
PR	02-JUL-1999;	99US-0142390P.
PR	08-JUL-1999;	99US-0142803P.
PR	09-JUL-1999;	99US-0142920P.
PR	12-JUL-1999;	99US-0142977P.
PR	13-JUL-1999;	99US-0143542P.
PR	14-JUL-1999;	99US-0143624P.
PR	15-JUL-1999;	99US-0144005P.
PR	16-JUL-1999;	99US-0144085P.
PR	16-JUL-1999;	99US-0144086P.
PR	19-JUL-1999;	99US-0144325P.
PR	19-JUL-1999;	99US-0144331P.
PR	19-JUL-1999;	99US-0144332P.
PR	19-JUL-1999;	99US-0144333P.
PR	19-JUL-1999;	99US-0144334P.
PR	19-JUL-1999;	99US-0144335P.
PR	20-JUL-1999;	99US-0144352P.
PR	20-JUL-1999;	99US-0144632P.
PR	20-JUL-1999;	99US-0144884P.
PR	21-JUL-1999;	99US-0144814P.
PR	21-JUL-1999;	99US-0145086P.
PR	21-JUL-1999;	99US-0145088P.
PR	21-JUL-1999;	99US-0145085P.
PR	22-JUL-1999;	99US-0145087P.
PR	22-JUL-1999;	99US-0145089P.
PR	22-JUL-1999;	99US-0145192P.
PR	23-JUL-1999;	99US-0145145P.
PR	23-JUL-1999;	99US-0145218P.
PR	23-JUL-1999;	99US-0145224P.
PR	26-JUL-1999;	99US-0145276P.
PR	27-JUL-1999;	99US-0145913P.
PR	27-JUL-1999;	99US-0145918P.
PR	27-JUL-1999;	99US-0145919P.
PR	28-JUL-1999;	99US-0145951P.
PR	02-AUG-1999;	99US-0146386P.
PR	02-AUG-1999;	99US-0146388P.
PR	02-AUG-1999;	99US-0146389P.
PR	03-AUG-1999;	99US-0147038P.
PR	04-AUG-1999;	99US-0147204P.
PR	04-AUG-1999;	99US-0147302P.
PR	05-AUG-1999;	99US-0147192P.
PR	05-AUG-1999;	99US-0147260P.
PR	06-AUG-1999;	99US-0147303P.
PR	06-AUG-1999;	99US-0147416P.
PR	09-AUG-1999;	99US-0147493P.
PR	09-AUG-1999;	99US-0147935P.
PR	10-AUG-1999;	99US-0148171P.
PR	11-AUG-1999;	99US-0148319P.
PR	12-AUG-1999;	99US-0148341P.
PR	13-AUG-1999;	99US-0148565P.
PR	13-AUG-1999;	99US-0148684P.
PR	16-AUG-1999;	99US-0149368P.
PR	17-AUG-1999;	99US-0149175P.
PR	18-AUG-1999;	99US-0149426P.
PR	20-AUG-1999;	99US-0149723P.
PR	20-AUG-1999;	99US-0149723P.
PR	20-AUG-1999;	99US-0149902P.
PR	23-AUG-1999;	99US-0149930P.
PR	23-AUG-1999;	99US-0150566P.
PR	26-AUG-1999;	99US-0150884P.
PR	27-AUG-1999;	99US-0151065P.
PR	27-AUG-1999;	99US-0151066P.
ER	27-AUG-1999;	99US-0151080P.
PR	30-AUG-1999;	99US-0151303P.
PR	31-AUG-1999;	99US-0151389P.
PR	01-SEP-1999;	99US-0151930P.
PR	07-SEP-1999;	99US-0152363P.
PR	10-SEP-1999;	99US-0153070P.
PR	13-SEP-1999;	99US-0153758P.

XX 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-01231180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125786P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132407P.
PR 05-MAY-1999; 99US-0132484P.
PR 06-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 11-MAY-1999; 99US-0132863P.
PR 14-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135112P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136382P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138009P.
PR 10-JUN-1999; 99US-0138504P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139452P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 24-JUN-1999; 99US-0140354P.
PR 28-JUN-1999; 99US-0140695P.
PR 29-JUN-1999; 99US-0140823P.
PR 30-JUN-1999; 99US-0140991P.
PR 01-JUL-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 02-JUL-1999; 99US-0142154P.
PR 06-JUL-1999; 99US-0142055P.
PR 08-JUL-1999; 99US-0142390P.
PR 09-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 15-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.

PR	23-APR-1999;	99US-01130891P.	99US-01448844P.
PR	28-APR-1999;	99US-01314499P.	99US-01448141P.
PR	30-APR-1999;	99US-01312048P.	99US-01450866P.
PR	30-APR-1999;	99US-01324077P.	99US-01450885P.
PR	04-MAY-1999;	99US-01324848P.	99US-01450885P.
PR	05-MAY-1999;	99US-01324856P.	99US-01450877P.
PR	06-MAY-1999;	99US-01324867P.	99US-01450899P.
PR	06-MAY-1999;	99US-01324878P.	99US-01451922P.
PR	07-MAY-1999;	99US-01328636P.	99US-01451451P.
PR	11-MAY-1999;	99US-01342566P.	99US-01452188P.
PR	14-MAY-1999;	99US-01342189P.	99US-01452242P.
PR	14-MAY-1999;	99US-01342198P.	99US-01452766P.
PR	14-MAY-1999;	99US-01342212P.	99US-01459133P.
PR	14-MAY-1999;	99US-01343707P.	99US-01459188P.
PR	18-MAY-1999;	99US-01347688P.	99US-01459198P.
PR	19-MAY-1999;	99US-01349411P.	99US-01459511P.
PR	20-MAY-1999;	99US-01351124P.	99US-01463866P.
PR	21-MAY-1999;	99US-01353533P.	99US-01463888P.
PR	24-MAY-1999;	99US-01356299P.	99US-01463889P.
PR	25-MAY-1999;	99US-01360212P.	99US-01463889P.
PR	27-MAY-1999;	99US-01363822P.	99US-01470388P.
PR	28-MAY-1999;	99US-01367822P.	99US-01472044P.
PR	01-JUN-1999;	99US-01372222P.	99US-01473022P.
PR	03-JUN-1999;	99US-01375288P.	99US-01471922P.
PR	04-JUN-1999;	99US-01375022P.	99US-01472606P.
PR	07-JUN-1999;	99US-0137724P.	99US-01473033P.
PR	08-JUN-1999;	99US-01380949P.	99US-01474166P.
PR	10-JUN-1999;	99US-01385406P.	99US-01474933P.
PR	10-JUN-1999;	99US-01388847P.	99US-01479335P.
PR	14-JUN-1999;	99US-01391199P.	99US-01481711P.
PR	16-JUN-1999;	99US-01394522P.	99US-01483119P.
PR	16-JUN-1999;	99US-01394533P.	99US-01483411P.
PR	17-JUN-1999;	99US-01394922P.	99US-01485655P.
PR	18-JUN-1999;	99US-01394549P.	99US-01486848P.
PR	18-JUN-1999;	99US-01394555P.	99US-01483688P.
PR	18-JUN-1999;	99US-01394566P.	99US-01491755P.
PR	18-JUN-1999;	99US-01394577P.	99US-01494266P.
PR	18-JUN-1999;	99US-01394588P.	99US-01497222P.
PR	18-JUN-1999;	99US-01394599P.	99US-01497233P.
PR	18-JUN-1999;	99US-01394606P.	99US-01499299P.
PR	18-JUN-1999;	99US-01394611P.	99US-01499022P.
PR	18-JUN-1999;	99US-01394622P.	99US-01499308P.
PR	18-JUN-1999;	99US-01394633P.	99US-01505666P.
PR	18-JUN-1999;	99US-01397505P.	99US-01508844P.
PR	18-JUN-1999;	99US-01397633P.	99US-01510666P.
PR	21-JUN-1999;	99US-01398177P.	99US-01510666P.
PR	22-JUN-1999;	99US-01398999P.	99US-01510808P.
PR	23-JUN-1999;	99US-01403533P.	99US-01513033P.
PR	23-JUN-1999;	99US-01403549P.	99US-01514388P.
PR	24-JUN-1999;	99US-01406955P.	99US-01519308P.
PR	28-JUN-1999;	99US-01408233P.	99US-01523633P.
PR	28-JUN-1999;	99US-01409911P.	99US-01530706P.
PR	30-JUN-1999;	99US-01412877P.	99US-01537588P.
PR	01-JUL-1999;	99US-01418422P.	99US-01540188P.
PR	02-JUL-1999;	99US-01421549P.	99US-01540399P.
PR	02-JUL-1999;	99US-01425555P.	99US-01547799P.
PR	08-JUL-1999;	99US-01423906P.	99US-01551339P.
PR	08-JUL-1999;	99US-01428033P.	99US-01554866P.
PR	09-JUL-1999;	99US-01429202P.	99US-01556599P.
PR	12-JUL-1999;	99US-01429777P.	99US-01564588P.
PR	13-JUL-1999;	99US-01435422P.	99US-01565966P.
PR	14-JUL-1999;	99US-01436249P.	99US-01571177P.
PR	15-JUL-1999;	99US-01440055P.	99US-01577533P.
PR	16-JUL-1999;	99US-01440855P.	99US-01578665P.
PR	16-JUL-1999;	99US-01440866P.	99US-01580299P.
PR	19-JUL-1999;	99US-01443255P.	99US-01582322P.
PR	19-JUL-1999;	99US-01443311P.	99US-01583699P.
PR	19-JUL-1999;	99US-01443322P.	99US-01592933P.
PR	1		

PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 63.0%; Score 51; DB 3; Length 289;
Best Local Similarity 69.2%; Pred. No. 30;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Oy 1 VLENKFPFPQET 13
Db 184 VTEMVPPPPQQT 196

Search completed: April 12, 2005, 08:18:25
Job time : 86 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 12, 2005, 07:49:30 ; Search time 17.5909 Seconds
(without alignments)
63.654 Million cell updates/sec

Title: US-10-691-330-7

Perfect score: 81

Sequence: 1 VLEMKFPPPPQETVT 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA.*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PTCUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	81	100.0	15	4	US-09-641-803-7
2	51	63.0	319	4	US-09-533-029-38
3	46	56.8	300	4	US-09-252-991A-29412
4	44	54.3	377	4	US-09-902-540-12610
5	44	54.3	498	4	US-09-949-016-9712
6	44	54.3	521	2	US-08-933-227-1
7	44	54.3	521	4	US-09-538-092-819
8	44	54.3	1523	4	US-09-538-092-955
9	43	53.1	498	4	US-09-949-016-11510
10	43	53.1	521	2	US-08-933-227-3
11	43	53.1	521	4	US-09-538-092-822
12	42	51.9	424	4	US-09-252-991A-20695
13	42	51.9	431	4	US-09-902-540-11153
14	42	51.9	568	4	US-09-248-796A-14918
15	42	51.9	613	4	US-09-949-016-9601
16	42	51.9	819	4	US-09-438-185A-1055
17	41	50.6	24	4	US-09-270-767-36510
18	41	50.6	24	4	US-09-270-767-51727
19	41	50.6	130	4	US-09-328-352-6607
20	41	50.6	274	4	US-09-902-540-15552
21	41	50.6	334	4	US-09-438-185A-11
22	41	50.6	1250	1	US-08-441-139-9
23	40.5	50.0	1070	4	US-09-961-403-3
24	40	49.4	82	4	US-09-621-976-4586
25	40	49.4	90	4	US-09-513-999C-6784
26	40	49.4	91	4	US-09-621-976-4585
27	40	49.4	366	3	US-09-176-657-2

28	40	49.4	366	4	US-09-421-299-2	Sequence 2, Appli
29	40	49.4	1217	4	US-09-949-016-7454	Sequence 7454, Ap
30	39	48.1	23	1	US-08-268-251-56	Sequence 56, Appl
31	39	48.1	23	5	PCT-US93-01112-56	Sequence 56, Appl
32	39	48.1	120	4	US-09-311-021-34	Sequence 34, Appl
33	39	48.1	153	4	US-09-252-991A-31659	Sequence 31659, A
34	39	48.1	252	4	US-09-248-796A-23055	Sequence 23055, A
35	39	48.1	259	4	US-09-248-796A-28192	Sequence 28192, A
36	39	48.1	282	4	US-09-270-767-39529	Sequence 39529, A
37	39	48.1	282	4	US-09-270-767-54746	Sequence 54746, A
38	39	48.1	320	4	US-09-248-796A-24758	Sequence 24758, A
39	39	48.1	403	4	US-09-270-767-37943	Sequence 37943, A
40	39	48.1	403	4	US-09-270-767-53160	Sequence 53160, A
41	39	48.1	425	4	US-09-270-767-45380	Sequence 45380, A
42	39	48.1	450	4	US-09-252-991A-26266	Sequence 26266, A
43	39	48.1	469	4	US-09-248-796A-14616	Sequence 14616, A
44	39	48.1	481	4	US-09-248-796A-14693	Sequence 14693, A
45	39	48.1	572	4	US-09-866-570B-65	Sequence 65, Appl

ALIGNMENTS

RESULT 1

US-09-641-803-7

; Sequence 7, Application US/09641803

; Patent No. 6500798

; GENERAL INFORMATION:

; APPLICANT: STANTON, G. John

; APPLICANT: HUGHES, Thomas K.

; APPLICANT: BOLDGOH, Istvan

; TITLE OF INVENTION: USE OF COLOSTRININ, CONSTITUENT PEPTIDES THEREOF, AND ANALOGS THEREOF AS OXIDATIVE STRESS REGULATORS

; FILE REFERENCE: 265.00220101

; CURRENT APPLICATION NUMBER: US/09/641.803

; CURRENT FILING DATE: 2000-08-17

; PRIOR FILING DATE: 1999-08-17

; NUMBER OF SEQ ID NOS: 34

; SOFTWARE: Patentin Ver. 2.1

; SEQ ID NO 7

; LENGTH: 15

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE: Description of Artificial Sequence: synthetic

; OTHER INFORMATION: peptide

; US-09-641-803-7

Query Match 100.0%; Score 81; DB 4; Length 15;

Best Local Similarity 100.0%; Pred. No. 4.le-06;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLEMKFPPPPQETVT 15

Db 1 VLEMKFPPPPQETVT 15

RESULT 2

US-09-533-029-38

; Sequence 38, Application US/09533029

; Patent No. 6664446

; GENERAL INFORMATION:

; APPLICANT: Heard, Jacqueline

; APPLICANT: Broun, Pierre

; APPLICANT: Riechmann, Jose-Luis

; APPLICANT: Keddie, James

; APPLICANT: Pineda, Omaira

; APPLICANT: Adam, Luc

; APPLICANT: Samaha, Raymond

; APPLICANT: Zhang, James

; APPLICANT: Yu, Guo-Liang

; APPLICANT: Ratcliffe, Oliver

us-10-691-330-7.ra1

Tue Apr 12 09:05:00 2005

```
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 12610
; LENGTH: 377
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-12610

Query Match      54.3%; Score 44; DB 4; Length 377;
Best Local Similarity 63.6%; Pred. No. 45;
Matches 7; Conservative 2; Mismatches 0; Gaps 0;

QY      2 LEMKPPPPQOE 12
      ||| |||||::
Db      341 LERHPPPPRQ 351

RESULT 5
US-09-949-016-9712 Application US/09949016.
; Sequence 9712.
; Patent No. 6812339.
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 9712
; LENGTH: 498
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-9712

Query Match      54.3%; Score 44; DB 4; Length 498;
Best Local Similarity 80.0%; Pred. No. 60;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      5 KFPFPPQETV 14
      ||| |||||
Db      208 KDFPPPMETV 217

RESULT 6
US-08-933-227-1
; Sequence 1. Application US/08933227
; Patent No. 5965394
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Guegler, Karl
; APPLICANT: Corley, Neil
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: HUMAN IMPORTIN ALPHA HOMOLOG
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS

; APPLICANT: Pilgrim, Marsha
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Reuber, Lynne
; TITLE OF INVENTION: DISEASE-INDUCED POLYNUCLEOTIDES
; FILE REFERENCE: MBI-010
; CURRENT APPLICATION NUMBER: US/09/533,029
; CURRENT FILING DATE: 2000-03-22
; EARLIER APPLICATION NUMBER: 60/125,814
; EARLIER FILING DATE: 1999-03-23
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 38
; LENGTH: 319
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: G502
US-09-533-029-38

Query Match      63.0%; Score 51; DB 4; Length 319;
Best Local Similarity 69.2%; Pred. No. 3.3;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      1 VLEMKPPPPQPT 13
      ||| |||||:
Db      184 VTEMVPPPPQPT 196

RESULT 3
US-09-252-991A-29412
; Sequence 29412. Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 29412
; LENGTH: 300
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-29412

Query Match      56.8%; Score 46; DB 4; Length 300;
Best Local Similarity 80.0%; Pred. No. 18;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 VLEMKPPPPP 10
      ||: |||||
Db      221 VVENKFPFDP 230

RESULT 4
US-09-902-540-12610
; Sequence 12610. Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
```


;; SOFTWARE: FastSeq for Windows Version 2.0
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/933,227
;; FILING DATE: Filed Herewith
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER:
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Billings, Lucy J.
;; REGISTRATION NUMBER: 36,749
;; REFERENCE/DOCKET NUMBER: PF-0394 US
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 650-855-0555
;; TELEFAX: 650-845-4166
;; INFORMATION FOR SEQ ID NO: 1:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 521 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; IMMEDIATE SOURCE:
;; LIBRARY: THPIPLB01
;; CLONE: 10403
US-08-933-227-1

Query Match 54.3%; Score 44; DB 2; Length 521;
Best Local Similarity 80.0%; Pred. No. 63;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 5 KFPFPPQETV 14
| | | | |
Db 231 KDPFPPMETV 240

RESULT 7
US-09-538-092-819
; Sequence 819, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Giot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CuratSeqFormatter Version 0.9
; SEQ ID NO 819
; LENGTH: 521
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Polypeptide Accession Number 000505
US-09-538-092-819

Query Match 54.3%; Score 44; DB 4; Length 521;
Best Local Similarity 80.0%; Pred. No. 63;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 5 KFPFPPQETV 14
| | | | |
Db 231 KDPFPPMETV 240

RESULT 8
US-09-538-092-955
; Sequence 955, Application US/09538092

;; Patent No. 6753314
;; GENERAL INFORMATION:
;; APPLICANT: Giot, Loic
;; APPLICANT: Mansfield, Traci A.
;; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
;; FILE REFERENCE: 15966-542
;; CURRENT APPLICATION NUMBER: US/09/538,092
;; CURRENT FILING DATE: 2000-03-29
;; PRIOR APPLICATION NUMBER: 60/127,352
;; PRIOR FILING DATE: 1999-04-01
;; PRIOR APPLICATION NUMBER: 60/178,965
;; PRIOR FILING DATE: 2000-02-01
;; NUMBER OF SEQ ID NOS: 1387
;; SOFTWARE: CuratSeqFormatter Version 0.9
;; SEQ ID NO 955
;; LENGTH: 1523
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: misc_feature
;; LOCATION: (0)...(0)
;; OTHER INFORMATION: Polypeptide Accession Number P18583
US-09-538-092-955

Query Match 54.3%; Score 44; DB 4; Length 1523;
Best Local Similarity 58.3%; Pred. No. 1.9e+02;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 3 EMKFPFPPQETV 14
| | | | |
Db 698 EKEVPPPKETL 709

RESULT 9
US-09-949-016-11510
; Sequence 11510, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11510
; LENGTH: 498
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-11510

Query Match 53.1%; Score 43; DB 4; Length 498;
Best Local Similarity 70.0%; Pred. No. 84;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 5 KFPFPPQETV 14
| | | | |
Db 208 KDPFPPMETI 217

RESULT 10
US-08-933-227-3
; Sequence 3, Application US/08933227
; Patent No. 5965394
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga

us-10-691-330-7.ra1

Tue Apr 12 09:05:00 2005

```

; APPLICANT: Guegler, Karl
; APPLICANT: Corley, Neil
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: HUMAN IMPORTIN ALPHA HOMOLOG
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/933,227
; FILING DATE: Filed Herewith
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0394 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 521 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 1928975
; US-08-933-227-3

Query Match 53.1%; Score 43; DB 2; Length 521;
Best Local Similarity 70.0%; Pred. No. 89;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 5 KPPPPQETV 14
DB 231 KPPPPMETI 240

RESULT 11
US-09-538-092-822
; Sequence 822, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Giot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CuratSeqFormat Version 0.9
; SEQ ID NO 822
; LENGTH: 521
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature

US-08-933-227-3
Query Match 53.1%; Score 43; DB 2; Length 521;
Best Local Similarity 70.0%; Pred. No. 89;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 5 KPPPPQETV 14
DB 231 KPPPPMETI 240

RESULT 11
US-09-538-092-822
; Sequence 822, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Giot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CuratSeqFormat Version 0.9
; SEQ ID NO 822
; LENGTH: 521
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature

; LOCATION: (0)...(0)
; OTHER INFORMATION: Polypeptide Accession Number 000629
US-09-538-092-822

Query Match 53.1%; Score 43; DB 4; Length 521;
Best Local Similarity 70.0%; Pred. No. 89;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 5 KPPPPQETV 14
DB 231 KPPPPMETI 240

RESULT 12
US-09-252-991A-20695
; Sequence 20695, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20695
; LENGTH: 424
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-20695

Query Match 51.9%; Score 42; DB 4; Length 424;
Best Local Similarity 77.8%; Pred. No. 1e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 7 PPPPPQETV 15
DB 34 PPPPPQETV 42

RESULT 13
US-09-902-540-11153
; Sequence 11153, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 11153
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
; US-09-902-540-11153

Query Match 51.9%; Score 42; DB 4; Length 431;
Best Local Similarity 70.0%; Pred. No. 1e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 5 KPPPPQETV 14
DB 366 KPPPPDEAV 375

```

RESULT 14

US-09-248-796A-14918
; Sequence 14918, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248.796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 14918
; LENGTH: 568
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-14918

Query Match 51.9%; Score 42; DB 4; Length 568;
Best Local Similarity 70.0%; Pred. No. 1.4e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 6 FPPPPQETVT 15
 |||||
Db 514 FPPPPPTTTT 523

RESULT 15

US-09-949-016-9601
; Sequence 9601, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949.016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9601
; LENGTH: 613
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-9601

Query Match 51.9%; Score 42; DB 4; Length 613;
Best Local Similarity 50.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLEMKFPPPP 10
 :::|||||
Db 3 LMELQYPPPP 12

Search completed: April 12, 2005, 08:07:53
Job time : 17.5909 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 12, 2005, 07:49:30 ; Search time 46.5 Seconds
(without alignments)
107.096 Million cell updates/sec

Title: US-10-691-330-7

Perfect score: 81

Sequence: 1 VLEMKFPPTQETVT 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1418010 seqs, 33197259 residues

Total number of hits satisfying chosen parameters: 1418010

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US10_NEW PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US11_NEW PUB.pep.*
- 19: /cgn2_6/ptodata/1/pubpaa/US60_NEW PUB.pep.*
- 20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	81	100.0	15	14 US-10-281-652-7	Sequence 7, Appli
2	81	100.0	15	17 US-10-691-330-7	Sequence 7, Appli
3	51	63.0	289	15 US-10-225-068-104	Sequence 104, App
4	51	63.0	289	15 US-10-374-780A-2014	Sequence 2014, Ap
5	51	63.0	289	15 US-10-412-699B-314	Sequence 314, App
6	51	63.0	289	17 US-10-495-918-140	Sequence 140, App
7	51	63.0	319	10 US-09-533-029-38	Sequence 38, Appl
8	51	63.0	319	14 US-10-286-264-22	Sequence 22, Appl
9	51	63.0	319	14 US-10-278-536-148	Sequence 148, App
10	47	58.0	179	16 US-10-767-701-56806	Sequence 56806, A
11	46	56.8	1196	15 US-10-369-493-3649	Sequence 3649, Ap
12	45	55.6	56	15 US-10-424-599-158309	Sequence 158309,
13	45	55.6	78	9 US-09-864-761-37299	Sequence 37299, A

14	54.3	61	15	US-10-424-599-220767	Sequence 220767,
15	54.3	95	16	US-10-437-963-127360	Sequence 127360,
16	54.3	102	15	US-10-424-599-170529	Sequence 170529,
17	54.3	111	16	US-10-437-963-176514	Sequence 176514,
18	54.3	137	15	US-10-424-599-265006	Sequence 265006,
19	54.3	158	15	US-10-424-599-161423	Sequence 161423,
20	54.3	166	15	US-10-225-066A-86	Sequence 86, Appl
21	54.3	166	15	US-10-225-067-10	Sequence 10, Appl
22	54.3	166	15	US-10-374-780A-2810	Sequence 2810, Ap
23	54.3	174	15	US-10-425-114-40629	Sequence 40629, A
24	54.3	289	14	US-10-011-585A-167	Sequence 167, App
25	54.3	361	9	US-09-864-761-33614	Sequence 33614, A
26	54.3	361	9	US-09-864-761-34141	Sequence 34141, A
27	54.3	420	16	US-10-437-963-106810	Sequence 106810,
28	54.3	515	16	US-10-437-963-169499	Sequence 169499,
29	54.3	521	15	US-10-231-956A-154	Sequence 154, App
30	54.3	529	9	US-09-925-301-866	Sequence 866, App
31	54.3	634	16	US-10-437-963-196028	Sequence 196028,
32	54.3	1047	14	US-10-029-386-33522	Sequence 33522, A
33	54.3	1087	14	US-10-029-386-32407	Sequence 32407, A
34	53.1	157	15	US-10-424-599-226767	Sequence 226767,
35	53.1	158	15	US-10-425-114-37363	Sequence 37363, A
36	53.1	176	16	US-10-437-963-113326	Sequence 113326,
37	53.1	201	16	US-10-437-963-103502	Sequence 103502,
38	53.1	214	15	US-10-424-599-240944	Sequence 240944,
39	53.1	219	9	US-09-925-301-1039	Sequence 1039, Ap
40	53.1	219	10	US-09-997-003-35	Sequence 35, Appl
41	53.1	219	10	US-09-997-003-47	Sequence 47, Appl
42	53.1	329	14	US-10-106-698-4702	Sequence 4702, Ap
43	53.1	337	16	US-10-437-963-181432	Sequence 181432,
44	53.1	414	15	US-10-424-599-213840	Sequence 213840,
45	53.1	417	16	US-10-437-963-143835	Sequence 143835,

ALIGNMENTS

RESULT 1

US-10-281-652-7
; Sequence 7, Application US/10281652
; Publication No. US20030091608A1
; GENERAL INFORMATION:
; APPLICANT: STANTON, G. John
; APPLICANT: HUGHES, Thomas K.
; APPLICANT: BOLDOGH, Istvan
; TITLE OF INVENTION: USE OF COLOSTRIN, CONSTITUENT PEPTIDES THEREOF, AND ANALOGS THEREOF AS OXIDATIVE STRESS REGULATORS
; FILE REFERENCE: 265.00220101
; CURRENT APPLICATION NUMBER: US/10/281.652
; PRIOR FILING DATE: 2002-10-28
; PRIOR FILING DATE: US/09/641,803
; PRIOR FILING DATE: 2000-08-17
; PRIOR APPLICATION NUMBER: 60/149,310
; PRIOR FILING DATE: 1999-08-17
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic peptide
; OTHER INFORMATION: peptide
US-10-281-652-7

Query Match 100.0%; Score 81; DB 14; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.00023;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLEMKFPPTQETVT 15

Db 1 VLEMKFPPTQETVT 15

us-10-691-330-7.rapb

Tue Apr 12 09:05:00 2005

```

; NAME/KEY: DOMAIN
; LOCATION: (10)...(155)
; OTHER INFORMATION: Conserved domain
US-10-225-068-104

Query Match      63.0%; Score 51; DB 15; Length 289;
Best Local Similarity 69.2%; Pred. No. 40;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VLEMKFPPPPQET 13
Db 184 VTEMVMPPPQET 196

RESULT 4
US-10-374-780A-2014
; Sequence 2014, Application US/10374780A
; Publication No. US20040019927A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, Bradley K
; APPLICANT: Riechmann, Jose Luis
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Heard, Jacqueline E
; APPLICANT: Haake, Volker
; APPLICANT: Creelman, Robert A
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Adam, Luc J
; APPLICANT: Reuber, T. Lynne
; APPLICANT: Reddie, James
; APPLICANT: Broun, Pierre E
; APPLICANT: Pilgrim, Marsha L
; APPLICANT: Dubell III, Arnold T
; APPLICANT: Pineda, Omaira
; APPLICANT: Yu, Guo-Liang
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES IN PLANTS
; FILE REFERENCE: MBI-0047 CIP US/10/374,780A
; CURRENT APPLICATION NUMBER: 09/837,944
; CURRENT FILING DATE: 2003-02-25
; PRIOR APPLICATION NUMBER: 60/310,847
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 09/934,455
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/336,049
; PRIOR FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/338,692
; PRIOR FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: 10/171,468
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 10/225,066
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 10/225,067
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 10/225,068
; PRIOR FILING DATE: 2002-08-09
; NUMBER OF SEQ ID NOS: 2906
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 2014
; LENGTH: 289
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: G502 Paralogous to G519
US-10-374-780A-2014

Query Match      63.0%; Score 51; DB 15; Length 289;
Best Local Similarity 69.2%; Pred. No. 40;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VLEMKFPPPPQET 13
Db 184 VTEMVMPPPQET 196

US-10-691-330-7
; Sequence 7, Application US/10691330
; Publication No. US20050042300A1
; GENERAL INFORMATION:
; APPLICANT: Boldogh, Istvan
; APPLICANT: Stanton, G. John
; APPLICANT: Georgiades, Jerzy A.
; APPLICANT: Hughes, Thomas K., Jr.
; APPLICANT: Kruzel, Marian
; TITLE OF INVENTION: USE OF COLOSTRININ, CONSTITUENT PEPTIDES THEREOF, AND ANALOGS
; TITLE OF INVENTION: THEREOF AS INHIBITORS OF APOPTOSIS AND OTHER CELLULAR DAMAGE
; FILE REFERENCE: 265.00390101
; CURRENT APPLICATION NUMBER: US/10/691,330
; CURRENT FILING DATE: 2003-10-22
; PRIOR APPLICATION NUMBER: US 60/420,369
; PRIOR FILING DATE: 2002-10-22
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 7
; LENGTH: 15
; TYPE: PRT
; ORGANISM: artificial
; FEATURE:
; OTHER INFORMATION: Synthetic Peptides
US-10-691-330-7

Query Match      100.0%; Score 81; DB 17; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.00023;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLEMKFPPPPQET 15
Db 1 VLEMKFPPPPQET 15

RESULT 3
US-10-225-068-104
; Sequence 104, Application US/10225068
; Publication No. US20030217383A1
; GENERAL INFORMATION:
; APPLICANT: Mendel Biotechnology, Inc.
; APPLICANT: Reuber, T. Lynne
; APPLICANT: Riechmann, Jose Luis
; APPLICANT: Heard, Jacqueline E.
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Adam, Luc J.
; APPLICANT: Dubell, Arnold T.
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Pineda, Omaira
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Broun, Pierre E.
; TITLE OF INVENTION: STRESS-RELATED POLYNUCLEOTIDES AND
; TITLE OF INVENTION: POLYPEPTIDES IN PLANTS
; FILE REFERENCE: 514442002040
; CURRENT APPLICATION NUMBER: US/10/225,068
; CURRENT FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 60/310,847
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/336,049
; PRIOR FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/338,692
; PRIOR FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: 10/171,468
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 246
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 104
; LENGTH: 289
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:

```

```
RESULT 5
US-10-412-699B-314
; Sequence 314, Application US/10412699B
; Publication No. US20040045049A1
; GENERAL INFORMATION:
; APPLICANT: Mendel Biotechnology, Inc.
; APPLICANT: Zhang, James
; APPLICANT: Fromm, Michael E.
; APPLICANT: Heard, Jacqueline E.
; APPLICANT: Riechmann, Jose Luis
; APPLICANT: Adam, Luc J.
; APPLICANT: Pierre E.
; APPLICANT: Pineda, Omaira
; APPLICANT: Reuber, T. Lynne
; APPLICANT: Keddie, James S.
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Samaha, Raymond R.
; APPLICANT: Pilgrim, Marsha L.
; APPLICANT: Creelman, Robert A.
; APPLICANT: Dubell, Arnold N.
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Kumimoto, Roderick
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: Polynucleotides and Polypeptides in Plants
; FILE REFERENCE: MBI-0048CIP
; CURRENT APPLICATION NUMBER: US/10/412,699B
; CURRENT FILING DATE: 2003-04-10
; PRIOR APPLICATION NUMBER: 09/394,519
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: 09/489,376
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: 09/506,720
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 09/533,030
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,392
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,029
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/532,591
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,648
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/713,994
; PRIOR FILING DATE: 2000-11-16
; PRIOR APPLICATION NUMBER: 09/819,142
; PRIOR FILING DATE: 2001-03-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2011
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 314
; LENGTH: 289
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: G502
US-10-412-699B-314

Query Match 63.0%; Score 51; DB 15; Length 289;
Best Local Similarity 69.2%; Pred. No. 40;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VLEMKFPPPPQET 13
||| |||||
Db 184 VTENVMPPPQQT 196

RESULT 6
US-10-495-918-140
; Sequence 140, Application US/10495918
; Publication No. US20050009187A1
; GENERAL INFORMATION:
; APPLICANT: RIKEN
; TITLE OF INVENTION: A STRESS-RESPONSIVE PROMOTER AND A GENE CODING FOR A STRESS-RESPONSIVE TRANSCRIPTION FACTOR
; FILE REFERENCE: RFH14-024T
; CURRENT APPLICATION NUMBER: US/10/495,918
; CURRENT FILING DATE: 2004-05-18
; PRIOR APPLICATION NUMBER: JP 2001-353038
; PRIOR FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: JP 2002-20329
; PRIOR FILING DATE: 2002-01-29
; NUMBER OF SEQ ID NOS: 166
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 140
; LENGTH: 289
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: G502
US-10-495-918-140

Query Match 63.0%; Score 51; DB 10; Length 319;
Best Local Similarity 69.2%; Pred. No. 45;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VLEMKFPPPPQET 13
||| |||||
Db 184 VTENVMPPPQQT 196

RESULT 8
US-10-533-029-38
; Sequence 38, Application US/09533029
; Publication No. US20030046723A1
; GENERAL INFORMATION:
; APPLICANT: Heard, Jacqueline
; APPLICANT: Broun, Pierre
; APPLICANT: Riechmann, Jose-Luis
; APPLICANT: Keddie, James
; APPLICANT: Pineda, Omaira
; APPLICANT: Adam, Luc
; APPLICANT: Samaha, Raymond
; APPLICANT: Zhang, James
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Pilgrim, Marsha
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Reuber, Lynne
; TITLE OF INVENTION: DISEASE-INDUCED POLYNUCLEOTIDES
; FILE REFERENCE: MBI-010
; CURRENT APPLICATION NUMBER: US/09/533,029
; CURRENT FILING DATE: 2000-03-22
; EARLIER APPLICATION NUMBER: 60/125,814
; EARLIER FILING DATE: 1999-03-23
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 38
; LENGTH: 319
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: G502
US-10-533-029-38

Query Match 63.0%; Score 51; DB 10; Length 319;
Best Local Similarity 69.2%; Pred. No. 45;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VLEMKFPPPPQET 13
||| |||||
Db 184 VTENVMPPPQQT 196

RESULT 8
```

us-10-691-330-7.rapb

Tue Apr 12 09:05:00 2005

```

US-10-278-536-148
Query Match      63.0%; Score 51; DB 14; Length 319;
Best Local Similarity 69.2%; Pred. No. 45;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VLEMKFPPPPQET 13
DB 184 VTEMVMPPPQQT 196

RESULT 10
US-10-767-701-56806
; Sequence 56806, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 56806
; LENGTH: 179
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: 30949255.pep
US-10-767-701-56806

Query Match      58.0%; Score 47; DB 16; Length 179;
Best Local Similarity 100.0%; Pred. No. 87;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 PPPPQETV 14
DB 70 PPPPQETV 77

RESULT 11
US-10-369-493-3649
; Sequence 3649, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 3649
; LENGTH: 1196
; TYPE: PRT
; ORGANISM: Neurospora crassa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(1196)
; OTHER INFORMATION: unsure at all Xaa locations
US-10-369-493-3649

Query Match      56.8%; Score 46; DB 15; Length 1196;
Best Local Similarity 57.1%; Pred. No. 7.5e+02;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

US-10-286-264-22
; Sequence 22, Application US/10286264
; Publication No. US20030093837A1
; GENERAL INFORMATION:
; APPLICANT: Keddle, James
; APPLICANT: Riechmann, Jose-Luis
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Zhang, James
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Pineda, Omaira
; APPLICANT: Heard, Jacqueline
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Adam, Luc
; APPLICANT: Broun, Pierre
; APPLICANT: Reuber, Lynne
; APPLICANT: Pilgrim, Marsha
; APPLICANT: Samaha, Raymond
; TITLE OF INVENTION: POLYNUCLEOTIDES FOR SEED TRAIT ALTERATION
; FILE REFERENCE: MBI-008
; CURRENT APPLICATION NUMBER: US/10/286,264
; CURRENT FILING DATE: 2002-11-01
; PRIOR APPLICATION NUMBER: 60/125,814
; PRIOR FILING DATE: 1999-03-23
; NUMBER OF SEQ ID NOS: 165
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 22
; LENGTH: 319
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: G502
US-10-286-264-22

Query Match      63.0%; Score 51; DB 14; Length 319;
Best Local Similarity 69.2%; Pred. No. 45;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VLEMKFPPPPQET 13
DB 184 VTEMVMPPPQQT 196

RESULT 9
US-10-278-536-148
; Sequence 148, Application US/10278536
; Publication No. US20030131386A1
; GENERAL INFORMATION:
; APPLICANT: Samaha, Raymond
; APPLICANT: Heard, Jacqueline
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Pineda, Omaira
; APPLICANT: Reuber, Lynne
; APPLICANT: Riechmann, Jose-Luis
; APPLICANT: Keddle, James
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Pilgrim, Marsha
; APPLICANT: Adam, Luc
; APPLICANT: Broun, Pierre
; TITLE OF INVENTION: STRESS-INDUCED POLYNUCLEOTIDES
; FILE REFERENCE: MBI-011
; CURRENT APPLICATION NUMBER: US/10/278,536
; CURRENT FILING DATE: 2002-10-22
; PRIOR APPLICATION NUMBER: 60/125,814
; PRIOR FILING DATE: 1999-03-23
; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 148
; LENGTH: 319
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: G502

```


Qy 2 LEMKFPPTPT 15
| : ||||| |
Db 1089 LSLFPPPTPT 1102

RESULT 12

US-10-424-599-158309
; Sequence 158309, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 158309
; LENGTH: 56
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_113972C.1.pep
US-10-424-599-158309

Query Match 55.6%; Score 45; DB 15; Length 56;
Best Local Similarity 58.3%; Pred. No. 52;
Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VLEMKFPPTPT 12
| : ||||| |
Db 16 VLKIKPPPTPT 27

RESULT 13

US-09-864-761-37299
; Sequence 37299, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecmika-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24363.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668

; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 37299
; LENGTH: 78
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC010906.2
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.9
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.4
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.4
; OTHER INFORMATION: EXPRESSED IN HEL100, SIGNAL = 2.5
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.9
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.9
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.6
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.8
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3
US-09-864-761-37299

Query Match 55.6%; Score 45; DB 9; Length 78;
Best Local Similarity 70.0%; Pred. No. 72;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 4 MKPPPTPT 13
| : ||||| |
Db 39 LKTPPTPT 48

RESULT 14

US-10-424-599-220767
; Sequence 220767, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 220767
; LENGTH: 61
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_41383C.1.pep
US-10-424-599-220767

Query Match 54.3%; Score 44; DB 15; Length 61;
Best Local Similarity 87.5%; Pred. No. 77;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 6 FPPPTPT 13
| : ||||| |

Tue Apr 12 09:05:00 2005

Db 14 FPPPPSET 21

```

RESULT 15
US-10-437-963-127360
; Sequence 127360, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 127360
; LENGTH: 95
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_2981C.1.pap
US-10-437-963-127360

```

```

Query Match      54.3%; Score 44; DB 16; Length 95;
Best Local Similarity 80.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

Oy 3 EMKPPPPPPQE 12
   | | | | |
Db 5 ERKPPPPPPQE 14

```

Search completed: April 12, 2005, 07:55:22
Job time : 47.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 12, 2005, 07:49:31 ; Search time 10.7727 Seconds
(without alignments)
133.973 Million cell updates/sec

Title: US-10-691-330-7

Perfect score: 81

Sequence: 1 VLEMKFPPPPQETVT 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_79.*

1: PIR1.*

2: PIR2.*

3: PIR3.*

4: PIR4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	51	63.0	229	2	ATAF1 protein - Ar
2	51	63.0	289	2	TIN6.12 protein -
3	49	60.5	180	2	hypothetical prote
4	46	56.8	148	2	hypothetical prote
5	46	56.8	272	2	hypothetical prote
6	44	54.3	204	2	protein F3M18.20
7	44	54.3	528	2	DNA-binding protei
8	44	54.3	1203	2	DNA-binding protei
9	43	53.1	196	2	hypothetical prote
10	43	53.1	330	2	mucin JUL10 - huma
11	43	53.1	521	2	mucin JUL7 - huma
12	43	53.1	543	2	mucin JUL7 - huma
13	43	53.1	599	2	hypothetical prote
14	43	53.1	610	2	abcl1 protein hemol
15	43	53.1	688	2	transcription fact
16	43	53.1	709	2	hypothetical prote
17	42	51.9	79	2	hypothetical prote
18	42	51.9	153	2	protein F18014.31
19	42	51.9	181	2	hypothetical prote
20	42	51.9	182	2	myofibrillar prote
21	42	51.9	291	2	B-cell surface ant
22	42	51.9	440	2	protein F18014.33
23	42	51.9	448	2	transcription regu
24	42	51.9	475	2	legumain (EC 3.4.2
25	42	51.9	811	2	hypothetical prote
26	42	51.9	913	2	probable ATP-depen
27	41	50.6	181	2	C29E4.9 protein -
28	41	50.6	213	2	heat shock protein
29	41	50.6	270	2	probable C2H2-type

probable Ap2 domain
hypothetical prote
hypothetical prote
probable legumain
cell division prot
hypothetical prote
protein-tyrosine k
SSD1 protein - yea
adenomatous polyo
kinase-like protei
protein-tyrosine k
calmodulin-like pr
DedA family protei
ribosomal protein
Ig heavy chain C r
transcription regu

ALIGNMENTS

RESULT 1

S37101

ATAF1 protein - Arabidopsis thaliana (fragment)

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004

C:Accession: S37101

R:Rueth, J.; Schweyen, R.; Hirt, H.

submitted to the EMBL Data Library, August 1993

A:Reference number: S37100

A:Accession: S37101

A:Molecule type: mRNA

A:Residues: 1-229 <RUE>

A:Cross-references: UNIPROT:Q39013; EMBL:X74755; NID:g398603; PID:e85756; PID:g1345506

C:Genetics:

A:Gene: ATAF1

Query Match 63.0%; Score 51; DB 2; Length 229;
Best Local Similarity 69.2%; Pred. No. 1.6;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VLEMKFPPPPQET 13
Db 124 VTEMVMPPPQQT 136

RESULT 2

E86148

TIN6.12 protein - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004

C:Accession: E86148

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

Chen, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.;

anssen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: E86148

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-289 <STO>

A:Cross-references: UNIPROT:Q9LQ85; GB:AE005172; NID:g8671840; PID:AAF78403.1; GSPDB:GN

C:Genetics:

A:Map position: 1

Query Match 63.0%; Score 51; DB 2; Length 289;
Best Local Similarity 69.2%; Pred. No. 2.1;

Tue Apr 12 09:05:00 2005

```

Matches      9;  Conservative      1;  Mismatches      3;  Indels      0;  Gaps      0;

QY      1  VLEMKFPPPPQET 13
      |||  |||||  ||
Db      184  VTEMVMPPPPPQET 196

RESULT 3
B84768
Hypothetical protein At2g35430 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C:Accession: E84768
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
Eisen, D.; Niernman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: E84768
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-180 <STO>
A:Cross-references: UNIPROT:O82297; GB:AE002093; NID:G3608145; PIDN:AAC36178.1; GSPDB:GN
C:Genetics:
A:Gene: At2g35430
A:Map position: 2

Query Match      60.5%; Score 49; DB 2; Length 180;
Best Local Similarity 58.8%; Pred. No. 2.5;
Matches 10; Conservative 2; Mismatches 1; Indels 4; Gaps 1;

QY      3  EMKFPFPPPP-----QETVT 15
      |||  |||||  |||||
Db      100  ELRLPFPFPPPPNQETVT 116

RESULT 4
F83542
Hypothetical protein PA0822 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C:Accession: F83542
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; B
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: F83542
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-148 <STO>
A:Cross-references: UNIPROT:Q915B7; GB:AE004517; GB:AE004091; NID:99946710; PIDN:AAG0421
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA0822

Query Match      56.8%; Score 46; DB 2; Length 148;
Best Local Similarity 80.0%; Pred. No. 5.7;
Matches 8; Conservative 1; Mismatches 1; Indels 1; Gaps 0;

QY      1  VLEMKFPPPPPP 10
      |||  |||||  ||
Db      69  VVEMKFPFPPDP 78

RESULT 5
S24375
Hypothetical protein 3 - Pseudomonas sp. (DSM 6898) plasmid pKB740
C:Species: Pseudomonas sp.
C>Date: 10-Mar-1994 #sequence_revision 10-Mar-1994 #text_change 09-Jul-2004
C:Accession: S24375

R:Altenschmidt, U.; Bokranz, M.; Fuchs, G.
Eur. J. Biochem. 207, 715-722, 1992
A:Title: Novel aerobic 2-aminobenzoate metabolism. Nucleotide sequence of the plasmid ca
P.
A:Reference number: S24374; MUID:923339462; PMID:1633822
A:Accession: S24375
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-272 <ALT>
A:Cross-references: UNIPROT:Q02304; EMBL:X66604; NID:945865; PIDN:CAA47169.1; PID:945867
A:Genome: plasmid

Query Match      56.8%; Score 46; DB 2; Length 272;
Best Local Similarity 58.3%; Pred. No. 11;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY      2  LEMKFPFPPPPQET 13
      |||  |||||  ||
Db      153  LRRRFPFPPPPDDT 164

RESULT 6
B86410
Protein F3M18.20 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C:Accession: B86410
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federpiet, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.P.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: B86410
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-204 <STO>
A:Cross-references: UNIPROT:Q9SGN4; GB:AE005172; NID:G6560756; PIDN:AAF16756.1; GSPDB:GN
C:Genetics:
A:Gene: F3M18.20
A:Map position: 1

Query Match      54.3%; Score 44; DB 2; Length 204;
Best Local Similarity 60.0%; Pred. No. 16;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY      2  LEMKFPFPPPPQ 11
      |||  |||||  ||
Db      194  LDLNFPFPPPE 203

RESULT 7
T52092
DNA-binding protein WRKY2 [imported] - common tobacco
C:Species: Nicotiana tabacum (common tobacco)
C>Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 09-Jul-2004
C:Accession: T52092
R:Wang, Z.; Yang, P.; Fan, B.; Chen, Z.
Plant J. 16, 515-522, 1998
A:Title: An oligo selection procedure for identification of sequence-specific DNA-bind
A:Reference number: Z25947; MUID:99097845; PMID:9881170
A:Accession: T52092
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-528 <WAN>
A:Cross-references: UNIPROT:Q9ZPL6; EMBL:AF096299; PIDN:AAD16139.1
C:Genetics:

```

A:Gene: WRKY2

Query Match 54.3%; Score 44; DB 2; Length 528;
Best Local Similarity 70.0%; Pred. No. 46;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 6 FPPPPQETVT 15
| | | | | : | | |
Db 79 FPPPPPTVS 88

RESULT 8

S26650
DNA-binding protein 5 - human
C:Species: Homo sapiens (man)
C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 01-Dec-2000
C:Accession: S26650; PNO099; I57397
R:Mattioni, T.; Hume, C.R.; Konigorski, S.; Hayes, P.; Osterweil, Z.; Lee, J.S.
Chromosome 101, 618-624, 1992
A:Title: A cDNA clone for a novel nuclear protein with DNA binding activity.
A:Reference number: S26650; MUID:93048367; PMID:1424986
A:Accession: S26650
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1203 <NAT>
A:Cross-references: EMBL:X63071
R:Bardichevsky, F.B.; Chumakov, I.M.; Kiselev, L.L.
Mol. Biol. (Mosk.) 22, 794-801, 1988
A:Title: Determination of the nucleotide sequence of the human genome son3 fragment: id
A:Reference number: PNO099; MUID:89039788; PMID:3054499
A:Accession: PNO099
A:Molecule type: mRNA
A:Residues: 'R', 574-699, 'PH', 702-1055 <BER>
A:Experimental source: embryo
A>Note: This paper is in Russian, with an English abstract
R:Bardichevsky, F.B.; Chumakov, I.M.; Kiselev, L.L.
Mol. Biol. 22, 639-646, 1988
A:Title: Determination of the nucleotide sequence of the son3 fragment of the human genc
A:Reference number: I57397
A:Accession: I57397
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 'R', 574-699, 'PH', 702-1055 <RES>
A:Cross-references: GB:M36428; NID:G338289; PIDN:AAA36624.1; PID:G338290
C:Genetics:
A:Introns: 790/3
C:Keywords: DNA binding
F:741-928/Region: arginine/serine-rich repeats

Query Match 54.3%; Score 44; DB 2; Length 1203;
Best Local Similarity 58.3%; Pred. No. 1.1e+02;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 3 EMKFPPTTQETV 14
| : | | | | : | | |
Db 585 EKEVPPPPKXTL 596

RESULT 9

T29602
hypothetical protein ZK816.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T29602
R:Nhan, M.; Le, T.
submitted to the EMBL Data Library, November 1995
A:Description: The sequence of C. elegans cosmid ZK816.
A:Reference number: Z20649
A:Accession: T29602
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-196 <NHA>
A:Cross-references: UNIPROT:Q23610; EMBL:U41018; PIDN:AAA82325.1; CESP:ZK816.4

C:Genetics:

A:Gene: CESP:ZK816.4
A:Introns: 40/3; 66/2; 105/2; 148/1
C:Superfamily: Caenorhabditis elegans hypothetical protein ZK816.4

Query Match 53.1%; Score 43; DB 2; Length 196;
Best Local Similarity 70.0%; Pred. No. 22;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 EMKFPPTTQETV 12
| : | | | | : | | |
Db 19 QFKFPPTTQETV 28

RESULT 10

S37595
mucin JUL10 - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
C:Accession: S37595; S35048
R:Aubert, J.
submitted to the EMBL Data Library, September 1993
A:Reference number: S37593
A:Accession: S37595
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-330 <AUB>
A:Cross-references: UNIPROT:Q14882; EMBL:X74956; NID:G407052; PIDN:CAA52911.1; PID:G4070
R:Dufoisse, J.; Porchet, N.; Audie, J.P.; Guyonnet Duperrat, V.; Laine, A.; van-Seueningen,
Biochem. J. 293, 329-337, 1993
A:Title: Degenerate 87-base-pair tandem repeats create hydrophilic/hydrophobic alternati
A:Reference number: S35047; MUID:93343858; PMID:7916618
A:Accession: S35048
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-330 <DUF>

Query Match 53.1%; Score 43; DB 2; Length 330;
Best Local Similarity 77.8%; Pred. No. 38;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 7 PPPPQETVT 15
| | | | | : | | | |
Db 186 PPPPQETLT 194

RESULT 11

JC5505
DNA helicase Q1 interacting protein 1 - human
N:Alternate names: nuclear localization signal receptor importin-alpha homolog
C:Species: Homo sapiens (man)
C>Date: 02-Sep-1997 #sequence_revision 05-Sep-1997 #text_change 09-Jul-2004
C:Accession: JC5505
R:Seki, T.; Tada, S.; Katada, T.; Enomoto, T.
Biochem. Biophys. Res. Commun. 234, 48-53, 1997
A:Title: Cloning of a cDNA encoding a novel importin-alpha homologue, Oip1: Discriminati
A:Reference number: JC5505; MUID:97312517; PMID:9168958
A:Accession: JC5505
A>Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-521 <SEK>
A:Cross-references: UNIPROT:O00629; DBJ:AB002533; NID:G1944124; PIDN:BAA19546.1; PID:G1
C:Comment: This protein interacts with the nuclear localization signal motif of SV40 T a
C:Superfamily: pendulin
F:11-51/Region: importin-beta binding #status predicted

Query Match 53.1%; Score 43; DB 2; Length 521;
Best Local Similarity 70.0%; Pred. No. 63;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 5 KFPPTTQETV 14
| | | | | : | | | |
Db 231 KDPPTTQETI 240

R;Wood, V.; Rajandream, M.A.; Barrell, B.G.; Taylor, K.; Harris, D.
submitted to the EMBL Data Library, September 1998
A:Reference number: Z21906
A:Accession: T40121
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-610 <WOO>
A:Cross-references: EMBL:AL031788; PIDN:CAA21176.1; GSPDB:GN00067; SPDB:SPBC2D10.18
A:Experimental source: strain 972h-; cosmid C2D10
C:Genetics:
A:Gene: abcl; SPBC2D10.18
A:Map position: 2

Query Match 53.1%; Score 43; DB 2; Length 610;
Best Local Similarity 46.2%; Pred. No. 75;
Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 VLEMKFPFPPQET 13
:|::: |||:|
DB 562 MLDLRLQPPPEET 574

RESULT 15
T09941
transcription factor VP1 - Craterostigma plantagineum
N:Alternate names: VP1 homolog
C:Species: Craterostigma plantagineum
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C:Accession: T09941
R;Chandler, J.W.; Bartels, D.
Mol. Gen. Genet. 256, 539-546, 1997
A:Title: Structure and function of the vp1 gene homologue from the resurrection plant Cr.
A:Reference number: Z16899; MUID:98075359; PMID:9413438
A:Accession: T09941
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-688 <CHA>
A:Cross-references: UNIPROT:O23774; EMBL:AJ000552; NID:g2288898; PIDN:CAA04184.1; PID:92
A:Experimental source: seed
C:Genetics:
A:Gene: vp1
C:Function:
A:Description: seed-specific transcriptional activator
C:Superfamily: rice transcription factor VP1
C:Keywords: DNA binding; seed; transcription factor; transcription regulation

Query Match 53.1%; Score 43; DB 2; Length 688;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 PPPPQET 13
:|:|:|:|
DB 321 PPPPQET 327

Search completed: April 12, 2005, 07:56:55
Job time : 11.7727 secs

RESULT 12
S35047
mucin JUL7 - human
C:Species: Homo sapiens (man)
C:Date: 10-Dec-1993 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
C:Accession: S35047
R;Dufosse, J.; Porchet, N.; Audie, J.P.; Guyonnet Duperrat, V.; Laine, A.; van-Seuningen, Biochem. J. 293, 329-337, 1993
A:Title: Degenerate 87-base-pair tandem repeats create hydrophilic/hydrophobic alternati
A:Reference number: S35047; MUID:93343858; PMID:7916618
A:Accession: S35047
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-543 <DUF>
A:Cross-references: UNIPROT:Q14879; EMBL:X74370; NID:g407081; PIDN:CAA52408.1; PID:g4070
A>Note: the authors translated the codon AAA for residue 63 as Gln and CGG for residue 3
A>Note: the amino acid sequence from Fig. 2 is inconsistent with the nucleotide sequence

Query Match 53.1%; Score 43; DB 2; Length 543;
Best Local Similarity 77.8%; Pred. No. 66;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 7 PPPPQETVT 15
:|:|:|:|
DB 361 PPPPQEQLT 369

RESULT 13
T39990
hypothetical protein SPBC25D12.02c - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: T39990
R;Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Whithead, S.; Chillingworth, T.; Churcher, submitted to the EMBL Data Library, July 1998
A:Reference number: Z21897
A:Accession: T39990
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-599 <LYN>
A:Cross-references: UNIPROT:O74354; EMBL:AL031158; PIDN:CAA20098.1; GSPDB:GN00067; SPDB:
A:Experimental source: strain 972h-; cosmid C25D12
C:Genetics:
A:Gene: SPDB:SPBC25D12.02c
A:Map position: 2
A:Introns: 48/2

Query Match 53.1%; Score 43; DB 2; Length 599;
Best Local Similarity 72.7%; Pred. No. 74;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 5 KFPFPPQETVT 15
:|:|:|:|
DB 415 KFPFPPQETLVT 425

RESULT 14
S71110
abc1 protein homolog precursor - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 28-Oct-1995 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C:Accession: S7111; T40121; S7110
R;Bonnetoy, N.; Kermorgant, M.; Brivet-Chevillotte, P.; Dujardin, G.
Mol. Gen. Genet. 251, 204-210, 1996
A:Title: Cloning by functional complementation, and inactivation, of the Schizosaccharom
A:Reference number: S7111; MUID:96242153; PMID:8668131
A:Accession: S7111
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 177-610 <BOW>
A:Cross-references: UNIPROT:Q92338; EMBL:X91616; NID:g1514640; PIDN:CAA62818.1; PID:e193

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 12, 2005, 07:49:31 ; Search time 71.1818 Seconds
(without alignments)
107.909 Million cell updates/sec

Title: US-10-691-330-7
Perfect score: 81
Sequence: 1 VLEMKFPPPPQETVT 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	55	67.9	1464	2 Q7SHT9	Q7sht9 neurospora
2	51	63.0	289	1 NAC2 ARATH	Q39013 arabidopsis
3	49	60.5	180	2 Q82297	Q82297 arabidopsis
4	47	58.0	375	2 Q7IFD7	Q7ifd7 mus musculus
5	47	58.0	375	2 Q99J35	Q99j35 mus musculus
6	47	58.0	526	2 Q92KS7	Q92ks7 rhizobium m
7	46	56.8	148	2 Q915B7	Q915b7 pseudomonas
8	46	56.8	272	2 Q02304	Q02304 pseudomonas
9	46	56.8	1499	2 Q7S266	Q7s266 neurospora
10	46	56.8	4587	2 Q9QXA3	Q9qxa3 mus musculus
11	45	55.6	114	2 Q9W026	Q9w026 drosophila
12	45	55.6	389	2 Q9F183	Q9f183 alcaligenes
13	45	55.6	803	1 ARH7 HUMAN	Q14155 homo sapien
14	44	54.3	166	2 Q9C5I3	Q9c5i3 arabidopsis
15	44	54.3	172	2 Q81YL7	Q81yl7 homo sapien
16	44	54.3	201	2 Q9GK20	Q9gk20 sus scrofa
17	44	54.3	204	2 Q9SGN4	Q9sgn4 arabidopsis
18	44	54.3	388	2 Q9NWJ6	Q9nwj6 homo sapien
19	44	54.3	337	2 Q9CT07	Q9ct07 mus musculus
20	44	54.3	420	2 Q6K8B2	Q6k8b2 oryza sativ
21	44	54.3	496	2 Q6TNT4	Q6tnt4 brachydanio
22	44	54.3	520	2 Q8JFT7	Q8jft7 brachydanio
23	44	54.3	520	2 Q7SXT4	Q7sxt4 brachydanio
24	44	54.3	521	1 IMA3 HUMAN	Q00505 homo sapien
25	44	54.3	521	1 IMA3 MOUSE	Q35344 mus musculus
26	44	54.3	521	2 Q81Y09	Q81yq9 homo sapien
27	44	54.3	521	2 Q61RB7	Q61rb7 xenopus lae
28	44	54.3	521	2 Q70PC6	Q70pc6 xenopus lae
29	44	54.3	528	2 Q92PL6	Q92pl6 nicotiana t
30	44	54.3	642	2 P87502	P87502 human adeno
31	44	54.3	642	2 Q6H1D2	Q6h1d2 human adeno

32	44	54.3	665	2 Q81WS8	Q81ws8 homo sapien
33	44	54.3	666	2 Q81Z27	Q81zz7 homo sapien
34	44	54.3	666	2 Q7Z5F6	Q7z5r6 homo sapien
35	44	54.3	1033	2 Q8T2L2	Q8t2l2 dictyostali
36	44	54.3	1365	2 Q6PKF0	Q6pke0 homo sapien
37	44	54.3	2426	1 SON HUMAN	P18583 homo sapien
38	43	53.1	61	2 Q91FX0	Q91fx0 chilo iride
39	43	53.1	68	2 Q6Q2Q3	Q6qzq3 felis silve
40	43	53.1	114	2 Q6Q2Q2	Q6qzq2 felis silve
41	43	53.1	166	2 Q8GK22	Q8gk22 sus scrofa
42	43	53.1	171	2 Q63401	Q63401 rattus norv
43	43	53.1	196	2 Q23610	Q23610 caenorhabdi
44	43	53.1	330	2 Q14882	Q14882 homo sapien
45	43	53.1	367	2 Q6AV33	Q6av33 oryza sativ

ALIGNMENTS

RESULT 1
Q7SHT9 PRELIMINARY; PRT; 1464 AA.
AC Q7SHT9; (TREMBlrel. 26, Created)
DT 01-MAR-2004 (TREMBlrel. 26, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Hypothetical protein.
GN Name=NCU02524.1;
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OR74A;
RA Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,
Jaffe D., FitzHugh W., Ma L.-J., Smirnov S., Purcell S., Rehm B.,
Elkins T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M.,
Qui D., Ianakiev P., Pedersen D., Nelson M., Washburne M.,
Selitrennikoff C.P., Kinsey J.A., Braun E.L., Zelter A., Schulte U.,
Kothe G.O., Jedd G., Mewes W., Staben C., Marcotte E., Greenberg D.,
Roy A., Foley K., Naylor J., Thomann N., Barrett R., Gnerre S.,
Kamal M., Kanvysselis M., Mauceli E., Bielke C., Rudd S., Frishman D.,
Krystofova S., Rasmussen C., Metzberg R.L., Perkins D.D., Kroken S.,
Cozoni C., Macino G., Catcheside D., Li W., Pratt R.J., Osmani S.A.,
Desouza C.C., Glass L., Orbach M.J., Berglund J., Voelker R.,
Yarden O., Piamann M., Sella S., Dunlap J., Radford A., Aramayo R.,
Natvig D.O., Alex L.A., Mannhaupt G., Eboole D.J., Freitag M.,
Paulsen I., Sachs M.S., Lander E.S., Nussbaum C., Birren B.,
"The Genome Sequence of the Filamentous Fungus Neurospora crassa.";
Nature 0:0-0(2003).
-!- CAUTION: The sequence shown here is derived from an
EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
preliminary data.
EMBL; AABX01000002; EAA36419.1; -.
HSSP; Q07960; 1AM4.
InterPro; IPR001849; PH.
InterPro; IPR011036; PH-related.
InterPro; IPR001098; RhoGAP.
InterPro; IPR008936; Rho_GAP.
Pfam; PF00169; PH; 1.
Pfam; PF00620; RhoGAP; 1.
PROSITE; PS50003; PH DOMAIN; 1.
PROSITE; PS50236; RHO GAP; 1.
KW Hypothetical protein.
SQ SEQUENCE 1464 AA; 159861 MW; 6EE3094BBAB0357 CRC64;

Query Match 67.9%; Score 55; DB 2; Length 1464;
Best Local Similarity 71.4%; Pred. No. 31;
Matches 10; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
Qy 2 LEMKFPFPPQETVT 15
|||:|||||:

Tue Apr 12 09:05:00 2005

```

Db      546 LEVMSPPPOETVS 559

RESULT 2
NAC2 ARATH STANDARD; PRT; 289 AA.
AC Q39013; OSLO85;
DT 25-OCT-2004 (Rel. 45, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE NAC-domain containing protein 2 (ANAC002)
GN Name=ATAF1; OrderedLocusName=At1G01720; ORFName=TFIN.12;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC euroids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Rueth J., Schweyen R., Hirt H.;
RL Submitted (AUG-1993) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=21016719; PubMed=11130712; DOI=10.1038/35048500;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altai H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
RA Dunn P., Egu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gill J.B., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziani A.,
RA Militscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
RA Uterback T., Van Aken S., Vayberg M., Vysotskaia V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
RT thaliana.";
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=22954850; PubMed=14593172; DOI=10.1126/science.1088305;
RA Yamada K., Lim J., Dale J.M., Chen H., Shinn P., Palm C.J.,
RA Southwick A.M., Wu H.C., Kim C.J., Nguyen M., Pham P.K., Cheuk R.F.,
RA Karlin-Newmann G., Liu S.X., Lam B., Sakano H., Wu T., Yu G.,
RA Miranda M., Quach H.L., Tripp M., Chang C.H., Lee J.M., Toriumi M.J.,
RA Chan M.M., Tang C.C., Onodera C.S., Deng J.M., Akiyama K., Ansari Y.,
RA Arakawa T., Banh J., Banno F., Bowser L., Brooks S.Y., Carninci P.,
RA Chao Q., Choy N., Enju A., Goldsmith A.D., Gurjal M., Hansen N.F.,
RA Hayashizaki Y., Johnson-Hopson C., Huan V.M., Iida K., Karnes M.,
RA Khan S., Koesema E., Ishida J., Jiang P.X., Jones T., Kawai J.,
RA Kamiya A., Meyers C., Nakajima M., Narusaka M., Seki M., Sakurai T.,
RA Satou M., Tamse R., Vayberg M., Wallender E.K., Wong C., Yamamura Y.,
RA Yuan S., Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
RT "Empirical analysis of transcriptional activity in the Arabidopsis
RT genome.";
RN [4]
RP Science 302:842-846(2003).
RN [5]
RP NOMBENCLATURE.
RX PubMed=15029955;
RA Ooka H., Satoh K., Doi K., Nagata T., Otomo Y., Murakami K.,
RA Matsubara K., Osato N., Kawai J., Carninci P., Hayashizaki Y.,
RA Suzuki K., Kojima K., Takahara Y., Yamamoto K., Kikuchi S.;
RT "Comprehensive analysis of NAC family genes in Oryza sativa and
RT Arabidopsis thaliana.";
RL DNA Res. 10:239-247(2003).
RN [6]
RP -!- DOMAIN: The NAC domain includes a DNA binding domain and a

```

```

CC dimerization domain.
CC -!- SIMILARITY: Contains 1 NAC domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X74755; CAA52771.1; ALT INIT.
CC EMBL; AC005273; AAE78403.1; -.
CC EMBL; AF370617; AAK43936.1; -.
CC PIR; E86148; E86148.
CC PIR; S37101; S37101.
CC GeneFarm; 4060; -.
CC InterPro; IPR003441; NAM.
CC Pfam; PF02365; NAM; 1.
CC PROSITE; PSS1005; NAC; 1.
CC DNA-binding; Nuclear protein; Transcription regulation.
FT DOMAIN 7 158
SQ SEQUENCE 289 AA; 32922 MW; DDAC62A101B6CE4B CRC64;

Query Match 63.0%; Score 51; DB 1; Length 289;
Best Local Similarity 69.2%; Pred. No. 20;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VLEMKFPPPPQET 13
Db 184 VTEMVMPPPPOOT 196

RESULT 3
OS2297 PRELIMINARY; PRT; 180 AA.
AC 082297
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein At2g35430.
DE Name=At2g35430.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC euroids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Rounsley S.D., Lin X., Kaul S., Shea T.P., Fujii C.Y., Mason T.M.,
RA Shen M., Ronning C.M., Fraser C.M., Somerville C.R., Venter J.C.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Town C.D., Kaul S.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
RL EMBL; AC005314; AAC36178.1; -.
RL PIR; E84768; E84768.
RL GO; GO:0003676; F:nucleic acid binding; IEA.
RL InterPro; IPR000571; Znf_CCH.
RL Pfam; PF00642; zf_CCH; 2.
RL SMART; SM00356; Znf_C3H1; 2.
KW Hypothetical protein.
SQ SEQUENCE 180 AA; 19991 MW; 9039E7F8120E4FF2 CRC64;

Query Match 60.5%; Score 49; DB 2; Length 180;
Best Local Similarity 58.8%; Pred. No. 23;
Matches 10; Conservative 2; Mismatches 1; Indels 4; Gaps 1;

QY 3 EMKFPFPPPP---QETVT 15
Db 100 ELRLPPPPPPPNQETVT 116

```



```

RESULT 4
Q71FD7 PRELIMINARY; PRT; 375 AA.
AC Q71FD7
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE CSX-associated LIM.
GN Name=Cal;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=1475752; DOI=10.1083/jcb.200309159;
RA Akazawa H., Kudoh S., Mochizuki N., Takekoshi N., Takano H., Nagai T.,
RA Komuro I.;
RT "A novel LIM protein Cal promotes cardiac differentiation by
RT association with CSX/NKX2-5."
RL J. Cell Biol. 164:395-405(2004).
CC -1- SIMILARITY: Contains 3 LIM zinc-binding domains.
DR EMBL; AF513359; AAQ08090.1; -.
DR HSSP; P32965; 1B8T.
DR InterPro; IPR000345; CytC_heme_BS.
DR InterPro; IPR001781; LIM.
DR Pfam; PF00412; LIM; 3.
DR ProDom; PD000094; LIM; 3.
DR SMART; SM00132; LIM; 3.
DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
DR PROSITE; PS00478; LIM_DOMAIN_1; 3.
DR PROSITE; PS50023; LIM_DOMAIN_2; 1.
KW LIM domain; Metal-binding; Zinc.
SQ SEQUENCE 375 AA; 41069 MW; 9BF3EDD16F0FA061 CRC64;

Query Match 58.0%; Score 47; DB 2; Length 375;
Best Local Similarity 61.5%; Pred. No. 99;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 3 EMKPPPPPPQETVT 15
Db 162 EEELPPPPPEPVT 174

RESULT 5
Q99J35 PRELIMINARY; PRT; 375 AA.
AC Q99J35
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE CSX-associated LIM.
GN Name=2410043F08Rik;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=NMRI; TISSUE=Mammary tumor; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Leitch S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.P., Capavanti T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S.J., Garcia A.M., Gay L.J., Rulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

```

```

RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=NMRI; TISSUE=Mammary tumor;
RA Strausberg R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 3 LIM zinc-binding domains.
DR EMBL; BC004777; AAH04777.1; -.
DR HSSP; P32965; 1B8T.
DR MGD; MGI:1921452; 2410043F08Rik.
DR InterPro; IPR000345; CytC_heme_BS.
DR InterPro; IPR001781; LIM.
DR Pfam; PF00412; LIM; 3.
DR ProDom; PD000094; LIM; 3.
DR SMART; SM00132; LIM; 3.
DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
DR PROSITE; PS00478; LIM_DOMAIN_1; 3.
DR PROSITE; PS50023; LIM_DOMAIN_2; 1.
KW LIM domain; Metal-binding; Zinc.
SQ SEQUENCE 375 AA; 41026 MW; 601DF99A8EB1BB80 CRC64;

Query Match 58.0%; Score 47; DB 2; Length 375;
Best Local Similarity 61.5%; Pred. No. 99;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 3 EMKPPPPPPQETVT 15
Db 162 EEELPPPPPEPVT 174

RESULT 6
Q92KS7 PRELIMINARY; PRT; 526 AA.
AC Q92KS7
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein SMC00336.
GN ORFNames=SMC00336;
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=1021;
RA MEDLINE=21396507; PubMed=11481430; DOI=10.1073/pnas.161294398;
RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,
RA Bolstad P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
RA Gohrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,
RA Pohl T., Portetelle D., Puhler A., Purnelle B., Rampersingh U.,
RA Renard C., Thebault P., Vandenbol M., Weidner S., Galibert F.;
RT "Analysis of the chromosome sequence of the legume symbiont
RT Sinorhizobium meliloti strain 1021."
RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
DR EMBL; AL591783; CAC41693.1; -.
DR GO; GO:0004497; F:monooxygenase activity; IEA.
DR GO; GO:0006725; P:aromatic compound metabolism; IEA.
DR InterPro; IPR003454; MmoB_DmpW.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 526 AA; 55373 MW; 6D657535288CB651 CRC64;

Query Match 58.0%; Score 47; DB 2; Length 526;
Best Local Similarity 75.0%; Pred. No. 1.4e+02;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

Tue Apr 12 09:05:00 2005

```

DR PROSITE; PS00527; RIBOSOMAL_S14; UNKNOWN_1.
KW Ligase; Monooxygenase.
SQ SEQUENCE 272 AA; 30631 MW; 17767C19A429050 CRC64;

Query Match 56.8%; Score 46; DB 2; Length 272;
Best Local Similarity 58.3%; Pred. No. 98;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 LEMKFPFPPPOET 13
ID | : ||||| : |
DB 68 LILKFPPLPOET 79

RESULT 7
QY15B7 PRELIMINARY; PRT; 148 AA.
AC Q915B7
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=PA0822;
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043; DOI=10.1038/35023079;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen."
RL Nature 406:959-964(2000).
DR EMBL; AE004517; AAG04211.1; -.
DR F01R; F83542; F83542.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 148 AA; 16617 MW; EB805603D1CBA4F CRC64;

Query Match 56.8%; Score 46; DB 2; Length 148;
Best Local Similarity 80.0%; Pred. No. 51;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLEMKFPFPP 10
ID | : ||||| : |
DB 69 VVEMKFPFDP 78

RESULT 8
QY2304 PRELIMINARY; PRT; 272 AA.
AC Q02304
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Plasmid DNA for 2-aminobenzoyl-CoA ligase, 2-aminobenzoyl-CoA
DE monooxygenase reductase and 5 open reading frames.
OS Pseudomonas sp.
OC Bacteria; Proteobacteria.
OX NCBI_TaxID=306;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=92339462; PubMed=1633822;
RA Altmerschmidt U., Bokranz M., Fuchs G.;
RT "Novel aerobic 2-aminobenzoyl-CoA ligase. Nucleotide sequence of the
RT plasmid carrying the gene for the flavoprotein 2-aminobenzoyl-CoA
RT monooxygenase/reductase in a denitrifying Pseudomonas sp.;"
RL Eur. J. Biochem. 207:715-722(1992).
DR EMBL; X66604; CAA47169.1; -.
DR F01R; S24375; S24375.
DR GO; GO:0016874; Ligase activity; IEA.
DR GO; GO:0004497; F:monooxygenase activity; IEA.
DR InterPro; IPR001209; Ribosomal_S14.
DR InterPro; IPR008900; Zet.
DR Pfam; PF05707; Zet; 1.

```

```

OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Whole embryo; PubMed=10741417;
RX MEDLINE=20203461; PubMed=10741417;
RA DOI=10.1002/(SICI)1097-0177(200003)217:3<233::AID-DVDY1>3.0.CO;2-O;
RA Cox B.T.M., Hadjantonakis A.K., Collins J., Magee A.I.;
RT "Cloning and expression throughout mouse development of mfat1 a
RT homologue of the Drosophila tumour suppressor gene fat.";
RL Dev. Dyn. 217:233-240(2000).
CC -1- SIMILARITY: Contains 33 cadherin domains.
DR EMBL: AJ250768; CAB65271.1; -.
DR HSSP: P08709; 1BF9.
DR MGB: MGI:109168; Fath.
DR GO: GO:0030054; C:cell junction; IDA.
DR GO: GO:0030175; C:filopodium; IDA.
DR GO: GO:0030027; C:lamellipodium; IDA.
DR GO: GO:0005515; P:protein binding; IPI.
DR GO: GO:0007015; P:actin filament organization; IMP.
DR GO: GO:0016337; P:cell-cell adhesion; IMP.
DR GO: GO:0007163; P:establishment and/or maintenance of cell po. .; IMP.
DR InterPro: IPR002126; Cadherin.
DR InterPro: IPR00985; ConA like_lec_gl.
DR InterPro: IPR00742; EGF_2.
DR InterPro: IPR001881; EGF_Ca.
DR InterPro: IPR006209; EGF_like.
DR InterPro: IPR009030; Grow_fac_recept.
DR InterPro: IPR001791; Laminin_G.
DR InterPro: IPR003129; TSP_N.
DR Pfam: PF00008; Cadherin; 33.
DR Pfam: PF00028; EGF; 3.
DR Pfam: PF02210; Laminin_G_2; 1.
DR PRINTS: PR00205; CADHERIN.
DR SMART: SM00112; CA; 33.
DR SMART: SM00179; EGF_Ca; 1.
DR SMART: SM00282; LamG; 1.
DR PROSITE: PS00232; CADHERIN_1; 17.
DR PROSITE: PS0268; CADHERIN_2; 33.
DR PROSITE: PS00022; EGF_1; UNKNOWN_2.
DR PROSITE: PS00026; EGF_3; 4.
DR PROSITE: PS00025; LAM_G_DOMAIN; 1.
KW Calcium; Calcium-binding.
FT NON_TER 4587 4587
SQ SEQUENCE 4587 AA; 506042 MW; 4D3F23BB05127CB4 CRC64;

Query Match 56.8%; Score 46; DB 2; Length 4587;
Best Local Similarity 63.6%; Pred. No. 2e+03;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Oy 2 LEMKPPPPPOE 12
Db 4430 IESDFPPPPPEE 4440
:| |||||:|
AC Q9W026 PRELIMINARY; PRT; 114 AA.
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE CG15859-PA.
GN ORFNames=CG15859;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazell R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkuch C., Baldwin D.,
RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Fertiera S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasok P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodger, Korley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426065; PubMed=12537568;
RA Celnik S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA George R.A., Hoskins R.A., Lavery T., Muzny D.M., Nelson C.R.,
RA Pacle J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT "Finishing a whole-genome shotgun: Release 3 of the Drosophila
RL melanogaster euchromatic genome sequence.";
RN Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RP [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426070; PubMed=12537573;
RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirskas R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celnik S.E.;
RT "The transposable elements of the Drosophila melanogaster euchromatin:
RL a genomic perspective.";
RN Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).
RP [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426069; PubMed=12537572;
RA Mira S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celnik S.E., de Grey A.D., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
RL systematic review.";
RN Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).

```

RA	Nagase T., Seki N., Tanaka A., Ishikawa K.-I., Nomura N.; "Prediction of the coding sequences of unidentified human genes. IV. The coding sequences of 40 new genes (KIAA0121-KIAA0160) deduced by analysis of cDNA clones from human cell line KG-1." DNA Res. 2:167-174(1995).
RN	[2]
RP	SEQUENCE FROM N.A. (ISOFORM 1).
RX	MEDLINE=98395067; PubMed=9726964; DOI=10.1074/jbc.273.37.23633; Bragrodia S., Taylor S.J., Jordan K.A., Van Aelst L., Cerione R.A.; "A novel regulator of p21-activated kinases." J. Biol. Chem. 273:23633-23636(1998).
RL	[3]
RN	SEQUENCE FROM N.A. (ISOFORM 3).
RP	TISSUE=Amygdala;
RC	Wambutt R., Heubner D., Mewes H.-W., Weil B., Wiemann S.; Submitted (JUL-2002) to the EMBL/GenBank/DDBJ databases.
RL	[4]
RN	SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 4).
RP	TISSUE=Brain, and Testis;
RC	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899; Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.F., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.M., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh P., Diatchenko L., Marusina K., Farmer A.A., Casavant T.L., Scheetz T.E., Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C., Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J., Bohak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalón D.K., Muzy D.M., Sodergren E.J., Lu X., Gibbs R.A., Whiting J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A., Fahney M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E., Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.; "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences."
RT	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RL	-!- FUNCTION: Acts as a Rac1 guanine nucleotide exchange factor (GEF) and can induce membrane ruffling [By similarity].
CC	-!- SUBUNIT: Interacts with Pak kinases through the SH3 domain. Interacts with GTP [By similarity].
CC	-!- ALTERNATIVE PRODUCTS:
CC	Event=Alternative splicing; Named isoforms=4;
CC	Name=4;
CC	IsoId=Q14155-4; Sequence=Displayed;
CC	Name=1;
CC	IsoId=Q14155-1; Sequence=VSP_011032, VSP_011035;
CC	Name=2;
CC	IsoId=Q14155-2; Sequence=VSP_011033;
CC	Name=3;
CC	IsoId=Q14155-3; Sequence=VSP_011034;
CC	-!- SIMILARITY: Contains 1 CH domain.
CC	-!- SIMILARITY: Contains 1 DBL-homology (DH) domain.
CC	-!- SIMILARITY: Contains 1 PH domain.
CC	-!- SIMILARITY: Contains 1 SH3 domain.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial/ entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@sib-sib.ch). -----
CC	EMBL; D63476; BAA09763.2; ALT_INIT.
DR	EMBL; ALG34228; CAD38906.1; ALT_INIT.
DR	EMBL; BC033905; AAH3905.1; --
DR	EMBL; BC050521; AAH5052.1; --
DR	EMBL; BC060776; AAH60776.1; --
DR	PDB; 1BY1; NMR; A=81-289.

DR Genew; HGNC:15607; ARHGGEF7.
 DR H-InVDB; HIX0011460; -.
 DR MIM; 605477; -.
 DR GO; GO:0007165; P:signal transduction; TAS.
 DR InterPro; IPR001331; GDS_CDC24.
 DR InterPro; IPR001849; PH.
 DR InterPro; IPR011036; PH related.
 DR InterPro; IPR000219; RHOGEF.
 DR InterPro; IPR001452; SH3.
 DR Pfam; PF00169; PH; 1.
 DR Pfam; PF00621; RHOGEF; 1.
 DR Pfam; PF00018; SH3; 1.
 DR PRINTS; PR00452; SH3DOMAIN.
 DR ProDom; PD000066; SH3; 1.
 DR SMART; SM00233; PH; 1.
 DR SMART; SM00325; RHOGEF; 1.
 DR SMART; SM00326; SH3; 1.
 DR PROSITE; PS50021; CH; 1.
 DR PROSITE; PS00741; DH_1; FALSE_NEG.
 DR PROSITE; PS50010; DH_2; 1.
 DR PROSITE; PS50003; PH_DOMAIN; 1.
 DR PROSITE; PS50002; SH3; 1.
 KW 3D-structure; Alternative splicing;
 KW Guanine-nucleotide releasing factor; SH3 domain.
 FT DOMAIN 1 132
 CH 132 SH3.
 FT DOMAIN 184 243
 DH 271 451
 FT DOMAIN 473 578
 FT VARSPLIC 1 178
 Missing (in isoform 1).
 FT VARSPLIC 56 105
 Missing (in isoform 2).
 FT VARSPLIC 85 105
 Missing (in isoform 3).
 FT VARSPLIC 733 803
 Missing (in isoform 1).
 FT VARSPLIC 343 343
 Missing (in isoform 1).
 FT VARSPLIC 552 552
 Missing (in isoform 1).
 FT VARSPLIC 636 636
 Missing (in isoform 1).
 FT VARSPLIC 740 740
 Missing (in isoform 1).
 SQ SEQUENCE 803 AA; 90011 MW; 613EBA839E6FDBD CRC64;
 Query Match 55.6%; Score 45; DB 1; Length 803;
 Best Local Similarity 62.5%; Pred. No. 4.4e+02;
 Matches 10; Conservative 1; Mismatches 3; Indels 2; Gaps 1;
 QY 2 LEMKPPP--PQETVT 15
 DB 83 LELLPPSPQPHLVT 98
 RESULT 14
 Q9C5I3 PRELIMINARY; PRT; 166 AA.
 ID Q9C5I3
 AC Q9C5I3;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Ethylene-responsive element binding factor, putative (ERF domain protein1).
 DE protein1).
 GN Name=At1g28370; Synonyms=ACEF1;
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,

RA Goldsmith A.D., Lee J.M., Quach H.L., Toriumi M., Yu G., Bowser L.,
 RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
 RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
 RA Miranda M., Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M.,
 RA Seki M., Shinn P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R.,
 RA Theologis A.;
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21380424; PubMed=11497705;
 RA Ohka M., Matsui K., Hiratsu K., Shiohara H., Ohme-Takagi M.;
 RT "Repression domains of class II ERF transcriptional repressors share
 an essential motif for active repression.";
 RL Plant Cell 13:1959-1968 (2001).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Takagi M.;
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Etgu P., Lee J.M.,
 RA Toriumi M., Yu G., Brooks S., Chao Q., Chen H., Karlin-Neumann G.,
 RA Kim C., Lam B., Miranda M., Nguyen M., Palm C.J., Shinn P.,
 RA Southwick A., Davis R.W., Ecker J.R., Theologis A.;
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22088475; PubMed=12093376;
 RA Haas B.J., Volfovsky N., Town C.D., Troukhan M., Alexandrov N.,
 RA Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;
 RT "Full-length messenger RNA sequences greatly improve genome
 annotation.";
 RL Genome Biol. 3:RESEARCH0029-RESEARCH0029 (2002).
 RN [6]
 RP SEQUENCE FROM N.A.
 RA Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
 RA Feldmann K.;
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF360232; AAK25942.1; -.
 DR EMBL; AB055882; BAB22911.1; -.
 DR EMBL; AY040032; AAK64090.1; -.
 DR EMBL; AY086783; AAM63833.1; -.
 DR HSSP; O80337; 2GCC.
 DR TRANSFAC; T05479; -.
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0003700; F:transcription factor activity; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro; IPR001471; TF ERF.
 DR PRINTS; PR00367; ETHRSPLEMT.
 DR ProDom; PD001423; TF ERF; 1.
 DR SMART; SM00380; AP2_1.
 SQ SEQUENCE 166 AA; 18612 MW; A327522DB20E5984 CRC64;
 Query Match 54.3%; Score 44; DB 2; Length 166;
 Best Local Similarity 60.0%; Pred. No. 1.1e+02;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 QY 2 LEMKPPPPPPQ 11
 DB 156 LDLNFPFPPPE 165
 RESULT 15
 Q8IYL7 PRELIMINARY; PRT; 172 AA:
 ID Q8IYL7
 AC Q8IYL7;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE APB11P protein.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

Tue Apr 12 09:05:00 2005

OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Blood;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RL and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Blood;
 RA Strausberg R.;
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC035636; AAH35636.1; -;
 SQ SEQUENCE 172 AA; 18685 MW; 21314883C4E8C7B9 CRC64;

Query Match 54.3%; Score 44; DB 2; Length 172;
 Best Local Similarity 46.7%; Pred. No. 1.2e+02;
 Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY . 1 VLEMKFPFPPPTVT 15
 ||:|
 Db 136 VLDLPLPFPFPLS 150

Search completed: April 12, 2005, 08:05:48
 Job time : 73.1818 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 12, 2005, 07:49:31 ; Search time 84 Seconds
(without alignments)
69.064 Million cell updates/sec

Title: US-10-691-330-8

Perfect score: 82

Sequence: 1 LKPPPKLKVEVFPFP 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq 16Dec04:*

- 1: geneseqp1980s:*
- 2: geneseqp1990s:*
- 3: geneseqp2000s:*
- 4: geneseqp2001s:*
- 5: geneseqp2002s:*
- 6: geneseqp2003as:*
- 7: geneseqp2003bs:*
- 8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	82	100.0	15	4	Aab72507 Colostrin
2	82	100.0	15	4	Aab59313 Ewe colos
3	82	100.0	15	4	Aab72253 Colostrin
4	82	100.0	15	4	Aab72539 Colostrin
5	82	100.0	15	5	Aao14584 Neural ce
6	82	100.0	15	5	Aam51043 Colostrin
7	82	100.0	15	5	Aae20235 Colostrin
8	82	100.0	15	8	Adn60302 Constitue
9	82	100.0	15	8	Adn74389 Ovine col
10	82	100.0	16	4	Aab59344 Ewe colos
11	47	57.3	436	7	Adn71271 Novel hum
12	47	57.3	867	7	Adn71288 Novel hum
13	44	53.7	650	8	Adp88214 Hansenula
14	43	52.4	175	4	Aag16487 Novel hum
15	43	52.4	240	3	Aag57452 Arabidops
16	43	52.4	246	3	Aag57451 Arabidops
17	43	52.4	249	8	Adn74103 Thale cre
18	43	52.4	954	8	Adr16190 Streptoco
19	43	52.4	1007	5	Abb91898 Herbicida
20	42	51.2	116	4	Aao03015 Human pol
21	42	51.2	184	6	Abm69539 Photorhab
22	42	51.2	232	7	Abp27516 Human pro
23	42	51.2	289	5	Abp27516 Streptoco
24	42	51.2	554	7	Adm03997 Human pro
25	42	51.2	562	4	Aau30319 Novel hum

26	42	51.2	950	7	Adn40022	Adn40022	Cancer/an
27	41.5	50.6	73	4	Aau31755	Aau31755	Novel hum
28	41.5	50.6	1675	5	Aau75109	Aau75109	Clathrin
29	41.5	50.6	1675	6	Ada50752	Ada50752	Human cla
30	41.5	50.6	1675	8	Abm81916	Abm81916	Tumour-as
31	41.5	50.6	1675	8	Adr69810	Adr69810	Wild-type
32	41	50.0	54	4	Aao04922	Aao04922	Human pol
33	41	50.0	74	4	Aag92276	Aag92276	C Glutami
34	41	50.0	77	4	Aam87301	Aam87301	Human imm
35	41	50.0	94	5	Abp03546	Abp03546	Human ORF
36	41	50.0	106	4	Aam06746	Aam06746	Human foe
37	41	50.0	106	4	Aam06758	Aam06758	Human foe
38	41	50.0	106	4	Aam06581	Aam06581	Human foe
39	41	50.0	119	4	Aao04769	Aao04769	Human pol
40	41	50.0	180	4	RAO01266	RAO01266	Human pol
41	41	50.0	188	4	AAU25583	AAU25583	Human G P
42	41	50.0	466	5	ABB93955	ABB93955	Herbicida
43	41	50.0	736	5	ABB55121	ABB55121	Lactococc
44	40.5	49.4	104	4	RAO06528	RAO06528	Human pol
45	40.5	49.4	170	8	Adn47517	Adn47517	Thermococ

ALIGNMENTS

RESULT 1
AAB72507
ID AAB72507 standard; peptide; 15 AA.
XX AC AAB72507;
XX AC
DT 09-MAY-2001 (first entry)
XX AC
DE Colostrinin peptide #8.
XX AC
KW Dermatological; oxidative stress regulator; colostrinin.
XX Unidentified.
OS
XX
FN WO200112650-A2.
XX
PD 22-FEB-2001.
XX
PF 17-AUG-2000; 2000WO-US022665.
XX
PR 17-AUG-1999; 99US-0149310P.
XX
(TEXA) UNIV TEXAS SYSTEM.
XX
Stanton GJ, Hughes TK, Boldogh I;
WPI; 2001-218342/22.
Modulating oxidative stress level in a cell, involves contacting the cell with an oxidative stress regulator selected from colostrinin, its constituent peptide, analog or their combinations.
Claim 6; Page 25; 48pp; English.

The present invention relates to a method for modulating the oxidative stress level in a cell or a patient, comprising contacting the cell with, or administering to the patient, an oxidative stress regulator selected from colostrinin, or its constituent peptide (e.g. the present peptide), to change the level of an oxidising species in the cell. The method can be used to treat oxidative damage to skin, by decreasing or preventing an increase in the level of damage to a biomolecule of the patient

SQ Sequence 15 AA;

Query Match 100.0%; Score 82; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 6.2e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

us-10-691-330-8.rag

Tue Apr 12 09:05:00 2005

KW dementia; neurodegenerative disease; Alzheimer's disease; psychosis;
 XX neurosis; infection.
 OS Synthetic.
 XX WO200111937-A2.
 XX 22-FEB-2001.
 XX 17-AUG-2000; 2000WO-US022818.
 XX 17-AUG-1999; 99US-0149311P.
 XX (TEXA) UNIV TEXAS SYSTEM.
 XX (REGE-) REGEN THERAPEUTICS PLC.
 XX Stanton GV, Hughes TK, Boldogh I, Georgiades J;
 XX WPI; 2001-202804/20.
 XX Inducing a cytokine and modulating an immune response, useful for
 PT treating central nervous system diseases and bacterial and viral
 PT infections, comprises administering colostrinin as an immunological
 PT regulator.
 XX Claim 1; Page 34; 50pp; English.
 XX Sequences AAB72246 - AAB72275 represent peptides derived from colostrinin,
 CC a proline rich polypeptide aggregate contained in colostrum. The peptides
 CC have immune response modulatory activity, and are capable of inducing
 CC cytokines. Colostrinin and its derived peptides are useful for inducing
 CC cytokine production, for modulating an immunological response and for
 CC inducing blood cell proliferation. The peptides are useful in the
 CC treatment of disorders of the central nervous system, neurological
 CC disorders, mental disorders, dementia, neurodegenerative diseases,
 CC Alzheimer's disease, motor neuron disease, psychosis, neurosis, chronic
 CC disorders of the immune system, bacterial and viral infections and
 CC acquired immunological deficiencies
 XX Sequence 15 AA;
 SQ Query Match 100.0%; Score 82; DB 4; Length 15;
 Best Local Similarity 100.0%; Pred. No. 6.2e-07; Indels 0; Gaps 0;
 Matches 15; Conservative 0; Mismatches 0;
 Qy 1 LKPPPKLKVEVFPFP 15
 Db 1 LKPPPKLKVEVFPFP 15
 RESULT 4
 AAB72539 standard; peptide; 15 AA.
 ID AAB72539
 XX AAB72539;
 AC 09-MAY-2001 (first entry)
 DT Colostrinin peptide #8.
 XX Neuroprotective; neural cell differentiation regulator; colostrinin;
 XX colostrum.
 XX Unidentified.
 XX WO200112651-A2.
 XX 22-FEB-2001.
 XX 17-AUG-2000; 2000WO-US022774.
 XX 17-AUG-1999; 99US-0149633P.
 XX

Qy 1 LKPPPKLKVEVFPFP 15
 Db 1 LKPPPKLKVEVFPFP 15
 RESULT 2
 AAB59313 standard; peptide; 15 AA.
 ID AAB59313
 XX AAB59313;
 AC 21-MAR-2001 (first entry)
 DT Ewe colostrinin peptide fragment A-4.
 XX Sheep; colostrinin; proline rich polypeptide; colostrum; immune disorder;
 XX central nervous system disorder; dietary supplement; beta-amyloid plaque.
 XX Ovis sp.
 XX WO200075173-A2.
 XX 14-DEC-2000.
 XX 02-JUN-2000; 2000WO-GB002128.
 XX 02-JUN-1999; 99GB-00012852.
 XX (REGE-) REGEN THERAPEUTICS PLC.
 XX Georgiades JA;
 XX WPI; 2001-071058/08.
 XX Peptides having an N-terminal amino acid sequence isolated from
 PT colostrinin for treating e.g. disorders of the central nervous system and
 PT immune system, viral and bacterial infections, and diseases characterized
 PT by amyloid plaques.
 XX Claim 7; Page 27; 63pp; English.
 XX The present invention provides the sequences of a number of peptides
 CC found in ewe's colostrinin. Colostrinin is the proline-rich polypeptide
 CC fragment of colostrum. These peptides can be used in the treatment of
 CC central nervous system disorders such as senile dementia, Parkinson's
 CC disease, Alzheimer's disease, psychosis and neurosis, immune system
 CC disorders such as bacterial and viral infections, to improve the
 CC development of a child's immune system, as a dietary supplement, and to
 CC promote the dissolution of beta-amyloid plaques
 XX Sequence 15 AA;
 SQ Query Match 100.0%; Score 82; DB 4; Length 15;
 Best Local Similarity 100.0%; Pred. No. 6.2e-07; Indels 0; Gaps 0;
 Matches 15; Conservative 0; Mismatches 0;
 Qy 1 LKPPPKLKVEVFPFP 15
 Db 1 LKPPPKLKVEVFPFP 15
 RESULT 3
 AAB72253 standard; peptide; 15 AA.
 ID AAB72253
 XX AAB72253;
 AC 14-MAY-2001 (first entry)
 DT Colostrinin derived cytokine inducing peptide SEQ ID 8.
 XX Colostrinin; immune response; cytokine; blood cell proliferation;
 XX central nervous system disorder; neurological disorder; mental disorder;
 KW

Tue Apr 12 09:05:00 2005

CC cytokine-inducing composition comprises a pharmaceutical carrier and an
CC active agent such as the present peptide. Cytokines induced by this
CC peptide in human leucocyte cultures include interferon-gamma, tumour
CC necrosis factor-alpha, interleukin-6 and interleukin-10
XX
SQ Sequence 15 AA;
Query Match 100.0%; Score 82; DB 5; Length 15;
Best Local Similarity 100.0%; Pred. No. 6.2e-07; Indels 0; Gaps 0;
Matches 15; Conservative 0; Mismatches 0;
QY 1 LKPPFKLKVEVFPFP 15
DB 1 LKPPFKLKVEVFPFP 15
RESULT 7
AAE20235
ID AAE20235 standard; peptide; 15 AA.
XX
AC AAE20235;
XX
DT 18-JUN-2002 (first entry)
DE Colostrinin constituent peptide #8.
XX Blood cell regulator; colostrinin; constituent peptide; oxidative stress;
KW therapy; oxidative damage; skin; aging; wound healing; cell replacement;
KW tissue; organ; cosmetic procedure; repair; regeneration; preservation;
KW transplantation; implantation; dermatological; vulnary.
XX
OS Unidentified.
XX
FH Key Location/Qualifiers
FT Modified-site 15 /note= "optionally C-terminal amide"
FT
XX
PN WO200213850-A1.
XX
PD 21-FEB-2002.
XX
PF 17-AUG-2000; 2000WO-US022776.
XX
PR 17-AUG-2000; 2000WO-US022776.
XX
PA (TEXA) UNIV TEXAS SYSTEM.
XX
PI Stanton GJ, Hughes TK, Boldogh I;
XX
DR WPI; 2002-269151/31.
XX
XX Composition useful for the modulation of blood cell proliferation in a
XX patient comprises a blood cell regulator selected from colostrinin, its
XX constituent peptide and/or analog.
XX
PS Claim 6; Page 25; 51pp; English.
XX
XX The invention relates to a composition which comprises a blood cell
XX regulator selected from colostrinin, its constituent peptide and/or
XX analogue. The invention is used for modulating the oxidative stress level
XX in a cell e.g. mammalian or human cell present in a cell culture, tissue,
XX organ, or organism; or for treating oxidative damage to the skin of a
XX patient e.g. animal or human; to modulate oxidative stress during/ after
XX a premature birth or normal birth, preventing/delaying aging in a
XX patient, enhancing wound healing, and the reduction of side effects of
XX cosmetic procedures. The method changes the level of an oxidizing species
XX in the cell, such as decreases or prevents increase in the level of
XX damage to a biomolecule of the patient selected from DNA, protein and/or
XX lipid, compared to the same conditions when the oxidative stress
XX regulator is not present. The modulation of oxidative stress results in
XX enhanced repair, regeneration, and replacement of cells, tissues and
XX organs (e.g. kidney, liver, pancreas, skin, and the other internal and
XX external organs), as well as enhanced preservation of such organs for

CC transplantation, implantation, or scientific research. The present
CC sequence is a colostrinin constituent peptide
XX
SQ Sequence 15 AA;
Query Match 100.0%; Score 82; DB 5; Length 15;
Best Local Similarity 100.0%; Pred. No. 6.2e-07; Indels 0; Gaps 0;
Matches 15; Conservative 0; Mismatches 0;
QY 1 LKPPFKLKVEVFPFP 15
DB 1 LKPPFKLKVEVFPFP 15
RESULT 8
ADN60302
ID ADN60302 standard; peptide; 15 AA.
XX
AC ADN60302;
XX
DT 29-JUL-2004 (first entry)
DE Constituent peptide of colostrinin SEQ ID NO:8.
XX
KW modulator; colostrinin; intracellular signaling molecule modulator;
KW 4HNE-mediated lipid peroxidation down regulation; apoptosis inhibition;
KW DNA damage; beta-amyloid; retinoic acid; cytostatic; 4HNE inhibitor;
KW 4HNE-protein adduct formation reduction;
KW 4HNE-mediated glutathione depletion inhibition; p53 inhibition;
KW c-Jun NH2-terminal kinase inhibition.
XX
OS Synthetic.
XX
PN WO2004037851-A2.
XX
PD 06-MAY-2004.
XX
PF 22-OCT-2003; 2003WO-US033423.
XX
PR 22-OCT-2002; 2002US-0420369P.
XX
PA (TEXA) UNIV TEXAS SYSTEM.
PA (BOLD/) BOLDOGH I.
PA (STAN/) STANTON J G.
PA (GEOR/) GEORGIADIS J A.
PA (HUGH/) HUGHES T K.
PA (KRUIZ/) KRUIZEL M.
XX
PI Boldogh I, Stanton JG, Georgiades JA, Hughes TK, Kruzel M;
XX
DR WPI; 2004-365494/34.
XX
XX Use of colostrinin for e.g. modulating an intracellular signaling
XX molecule in a cell, down regulating 4HNE-mediated lipid peroxidation in a
XX cell, inhibiting apoptosis in a cell, or protecting against DNA damage in
XX a cell.
XX
PS Claim 6; SEQ ID NO 8; 46pp; English.
XX
XX The present invention describes the use of a modulator selected from
XX colostrinin, its constituent peptide, its active analogue, and a
XX combination of these, for modulating an intracellular signaling molecule
XX in a cell, down regulating 4HNE-mediated lipid peroxidation in a cell,
XX inhibiting apoptosis in a cell, protecting against DNA damage in a cell,
XX and reducing the toxic effect of beta-amyloid or retinoic acid on a cell.
XX The modulator has cytostatic activity, and can be used as a 4HNE
XX inhibitor. The modulator is useful in the manufacture of a medicament for
XX reducing 4HNE-protein adduct formation, inhibiting 4HNE-mediated
XX glutathione depletion, inhibiting 4HNE-induced activation of p53 protein,
XX and/or inhibiting 4HNE-induced activation of c-Jun NH2-terminal kinases.
XX Colostrinin, or its constituent peptide or active analogue is useful for
XX inhibiting apoptosis in a cell, protecting against DNA damage in a cell,
XX or reducing the toxic effect of beta-amyloid or retinoic acid on a cell.

CC The present sequence represents a synthetic constituent peptide of
 CC colostrinin, which can be used as a modulator in the present invention.
 XX
 SQ Sequence 15 AA;

Query Match 100.0%; Score 82; DB 8; Length 15;
 Best Local Similarity 100.0%; Pred. No. 6.2e-07;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LKPPPKLKVEVFPP 15
 |||||
 Db 1 LKPPPKLKVEVFPP 15

RESULT 9

ADS74389
 ID ADS74389 standard; peptide; 15 AA.

XX
 AC ADS74389;

XX 16-DEC-2004 (first entry)

XX Ovine colostrinin peptide.

XX Colostrum; colostrinin; sheep; peptide purification.

XX Ovis aries.

XX WO2004081038-A1.

XX 23-SEP-2004.

XX 10-MAR-2004; 2004WO-GB001014.

XX 11-MAR-2003; 2003GB-00005552.

XX 08-MAR-2004; 2004GB-00005190.

XX (REG-) REGEN THERAPEUTICS PLC.

XX Georgiades JA, Polanowski A, Wilusz T, Kruzel ML;

XX WPI; 2004-677519/66.

XX Recovering peptides such as colostrinin from mammalian colostrum, by
 PT mixing colostrum with alcohol to form alcohol phase containing peptides
 PT and precipitate, separating alcohol phase from precipitate, and
 PT recovering alcohol phase.

XX Disclosure; SEQ ID NO 4; 41pp; English.

XX The present sequence is that of a peptide that can be recovered from
 CC ovine colostrinin using the method of the invention. The invention
 CC provides a method for the recovery of peptides (especially colostrinin)
 CC from colostrum in substantially pure, biologically active form and in
 CC high yield. The method involves mixing the colostrum with an alcohol to
 CC form an alcohol phase containing the colostrinin and a precipitate
 CC containing higher molecular weight caseins and other proteins. Best
 CC results are obtained using methanol or ethanol of at least 80%, and
 CC preferably up to 100%, purity. The alcohol phase is then separated from
 CC the precipitate, and the colostrinin is separated from the alcohol,
 CC preferably by evaporation, to form a colostrinin-rich phase, which is
 CC recovered. A precipitation agent, such as ammonium sulfate, may be added
 CC either to the alcohol phase or, preferably, to the colostrinin-rich phase
 CC to induce precipitation of the colostrinin peptides. The method is
 CC generally applicable to the separation of peptides from fluids containing
 CC higher molecular weight proteins, lipids, carbohydrates and/or nucleic
 CC acids.

XX Sequence 15 AA;

Query Match 100.0%; Score 82; DB 8; Length 15;
 Best Local Similarity 100.0%; Pred. No. 6.2e-07;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LKPPPKLKVEVFPP 15
 |||||
 Db 1 LKPPPKLKVEVFPP 15

RESULT 10

AAB59344
 ID AAB59344 standard; peptide; 16 AA.

XX
 AC AAB59344;

XX 21-MAR-2001 (first entry)

XX Ewe colostrinin peptide fragment derived sequence #4.

XX Sheep; colostrinin; proline rich polypeptide; colostrum; immune disorder;
 KW central nervous system disorder; dietary supplement; beta-amyloid plaque.

XX Ovis sp.

XX WO2000075173-A2.

XX 14-DEC-2000.

XX 02-JUN-2000; 2000WO-GB002128.

XX 02-JUN-1999; 99GB-00012852.

XX (REG-) REGEN THERAPEUTICS PLC.

XX Georgiades JA;

XX WPI; 2001-071058/08.

XX Peptides having an N-terminal amino acid sequence isolated from
 PT colostrinin for treating e.g. disorders of the central nervous system and
 PT immune system, viral and bacterial infections, and diseases characterized
 PT by amyloid plaques.

XX Claim 8; Page 27; 63pp; English.

XX The present invention provides the sequences of a number of peptides
 CC found in ewe's colostrinin. Colostrinin is the proline-rich polypeptide
 CC fragment of colostrum. These peptides can be used in the treatment of
 CC central nervous system disorders such as senile dementia, Parkinson's
 CC disease, Alzheimer's disease, psychosis and neurosis, immune system
 CC disorders such as bacterial and viral infections, to improve the
 CC development of a child's immune system, as a dietary supplement, and to
 CC promote the dissolution of beta-amyloid plaques

XX Sequence 16 AA;

Query Match 100.0%; Score 82; DB 4; Length 16;
 Best Local Similarity 100.0%; Pred. No. 6.7e-07;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LKPPPKLKVEVFPP 15
 |||||
 Db 2 LKPPPKLKVEVFPP 16

RESULT 11

ADE71271
 ID ADE71271 standard; protein; 436 AA.

XX
 AC ADE71271;

XX 29-JAN-2004 (first entry)

XX Novel human protein #25.

XX human; novel protein; drug.

XX Homo sapiens. 57.3%; Score 47; DB 7; Length 436;
 XX Best Local Similarity 61.5%; Pred. No. 20; Indels 0; Gaps 0;
 XX Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 XX JP2002345493-A.
 XX 03-DEC-2002.
 XX 29-MAR-2001; 2002JP-00049046.
 XX 29-MAR-2001; 2001JP-00095524.
 XX (KAZU-) ZH KAZUSA DNA KENKYUSHO.
 XX WPI; 2003-460885/44.
 XX N-PSDB; ADE71209.
 XX A gene and a protein encoded by it, used in drugs.
 XX Disclosure; Page 126-128; 257pp; Japanese.
 XX The invention comprises the amino acid and coding sequences of novel
 XX human proteins. The DNA and protein sequences of the invention are used
 XX in drugs. The present amino acid sequence represents a novel human
 XX protein of the invention.
 XX Sequence 436 AA;
 SQ

Query Match 57.3%; Score 47; DB 7; Length 436;
 Best Local Similarity 61.5%; Pred. No. 20;
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 1 LKPPPKLKVEVFP 13
 | | | | | : | |
 358 LTPYPKLKTALFP 370

QY 1 LKPPPKLKVEVFP 13
 Db 358 LTPYPKLKTALFP 370

RESULT 12
 ADE71288
 ID ADE71288 standard; protein; 867 AA.
 XX
 AC ADE71288;
 XX
 DT 29-JAN-2004 (first entry)
 XX
 DE Novel human protein #42.
 XX
 KW human; novel protein; drug.
 XX
 OS Homo sapiens.
 XX
 XX JP2002345493-A.
 XX
 XX 03-DEC-2002.
 XX
 XX 29-MAR-2001; 2002JP-00049046.
 XX
 XX 29-MAR-2001; 2001JP-00095524.
 XX
 XX (KAZU-) ZH KAZUSA DNA KENKYUSHO.
 XX
 XX WPI; 2003-460885/44.
 XX
 XX N-PSDB; ADE71226.
 XX
 XX A gene and a protein encoded by it, used in drugs.
 XX
 XX Disclosure; Page 189-192; 257pp; Japanese.
 XX
 XX The invention comprises the amino acid and coding sequences of novel
 XX human proteins. The DNA and protein sequences of the invention are used
 XX in drugs. The present amino acid sequence represents a novel human
 XX protein of the invention.
 XX Sequence 867 AA;
 SQ

Query Match 57.3%; Score 47; DB 7; Length 867;
 Best Local Similarity 61.5%; Pred. No. 42;
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 1 LKPPPKLKVEVFP 13
 | | | | | : | |
 114 LTPYPKLKTALFP 126

QY 1 LKPPPKLKVEVFP 13
 Db 114 LTPYPKLKTALFP 126

RESULT 13
 ADP88214
 ID ADP88214 standard; protein; 650 AA.
 XX
 AC ADP88214;
 XX
 DT 09-SEP-2004 (first entry)
 XX
 DE Hansenula polymorpha gamma-glutamyl cysteine synthetase HPGSH2.
 XX
 KW Gamma-glutamyl cysteine synthetase; HPGSH2; methanol-assimilating yeast;
 KW Hansenula polymorpha; glutathione biosynthesis; detoxification;
 KW heavy metal; adsorption; detection; mass production; liver function;
 KW food additive; nutritional supplement; adsorption; enzyme production;
 KW cell growth; enzyme.
 XX
 OS Pichia angusta.
 XX
 FH Key Location/Qualifiers
 FT Active-site 251..261
 FT Binding-site /label= MGFGMGXXCLO_motif
 FT /label= GWRVEFRPME motif
 FT /note= "Binds glutathione. Corresponds to SEQ ID NO:5"
 FT Binding-site 503
 FT /note= "Binds glutathione"
 XX
 XX JP2004173503-A.
 XX
 XX 24-JUN-2004.
 XX
 XX 22-NOV-2002; 2002JP-00339824.
 XX
 XX 22-NOV-2002; 2002JP-00339824.
 XX
 XX (KORE-) KOREA RES INST BIOSCIENCE & BIOTECHNOLOG.
 XX (CELL-) INST CELL BIOLOGY.
 XX (REAL-) REAL BIOTECH CO LTD.
 XX (KORI-) KORITSU KK.
 XX
 XX WPI; 2004-456234/43.
 XX
 XX N-PSDB; ADP88212, ADP88215.
 XX
 XX Novel genomic DNA comprising gamma-glutamyl cysteine synthetase gene and
 XX its promoter derived from Hansenula polymorpha, useful in mass production
 XX of glutathione.
 XX
 XX Claim 3; SEQ ID NO 3; 23pp; Japanese.
 XX
 XX The invention relates to a gamma-glutamyl cysteine synthetase named
 XX HPGSH2 (ADP88214) from the methanol-assimilating yeast Hansenula
 XX polymorpha (also known as Pichia angusta), nucleic acid sequences
 XX encoding HPGSH2 (ADP88212 and ADP88215), and the HPGSH2 promoter. The
 XX invention also relates to expression vectors and host Escherichia coli
 XX cells comprising the HPGSH2 coding sequence. HPGSH2 has 53% identity and
 XX 69% similarity to the Candida albicans gamma-glutamyl cysteine synthetase
 XX Gcsip, 47% identity and 61% similarity to the Saccharomyces cerevisiae
 XX ScGshp protein, 46% identity and 63% similarity to the
 XX Schizosaccharomyces pombe SpGcsip protein, and 45% identity and 63%
 XX similarity to the human HsGCSH protein. HPGSH2, like other gamma-glutamyl
 XX cysteine synthetases, is involved in the biosynthesis of glutathione,
 XX which plays a key role in the detoxification of substances such as heavy
 XX metals. HPGSH2-encoding nucleic acids may be used for generating
 XX

transformed host cells useful in the mass production of glutathione.
 CC Glutathione may be used to improve liver function, as a nutritional
 CC supplement or food additive (e.g., in sports drinks), as a detoxifying
 CC agent, as an agent for adsorption of heavy metals, and in enzyme
 CC production. The HpgSH2-encoding nucleic acids may also be used to restore
 CC cell growth function to glutathione synthesis-defective host cells. The
 CC HpgSH2 promoter may be used in reporter constructs useful in the
 CC detection of toxic substances such as heavy metals. The present sequence
 CC represents the Haneenuia polymorpha gamma-glutamyl cysteine synthetase
 CC HpgSH2.
 XX

SQ Sequence 650 AA;

Query Match 53.7%; Score 44; DB 8; Length 650;
 Best Local Similarity 46.7%; Pred. No. 98;
 Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 1 LKPPPKLKVEVFPF 15
 | | | | | | | | | |
 Db 148 LTSFPRMGVEIFTYP 162

RESULT 14

ID ABG16487 standard; protein; 175 AA.

AC ABG16487;

DT 18-FEB-2002 (first entry)

DE Novel human diagnostic protein #16478.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US008631.

XX 31-MAR-2000; 2000US-00540217.

PR 23-AUG-2000; 2000US-00649167.

XX (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

DR WPI; 2001-639362/73.

DR N-PSDB; AAS80674.

XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.

FS Claim 20; SEQ ID NO 46846; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (III) and its binding partners are useful in medical imaging
 CC of sites expressing (I). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in

CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
 CC amino acid sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 175 AA;

Query Match 52.4%; Score 43; DB 4; Length 175;
 Best Local Similarity 61.5%; Pred. No. 35;
 Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 LKPPPKLKVEVFP 13
 | | | | | | | | | |
 Db 27 LPPFPPLKFFIFP 39

RESULT 15

AAG57452
 ID AAG57452 standard; protein; 240 AA.

XX AAG57452;

DT 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 74041.

KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.

OS Arabidopsis thaliana.

XX EPI033405-A2.

PD 06-SEP-2000.

XX 25-FEB-2000; 2000EP-00301439.

XX 25-FEB-1999; 99US-0121825P.

PR 05-MAR-1999; 99US-0123180P.

PR 09-MAR-1999; 99US-0123548P.

PR 23-MAR-1999; 99US-0125788P.

PR 25-MAR-1999; 99US-0126284P.

PR 29-MAR-1999; 99US-0126785P.

PR 01-APR-1999; 99US-0127462P.

PR 06-APR-1999; 99US-0128234P.

PR 16-APR-1999; 99US-0129845P.

PR 19-APR-1999; 99US-0130077P.

PR 21-APR-1999; 99US-0130449P.

PR 23-APR-1999; 99US-0130510P.

PR 23-APR-1999; 99US-0130891P.

PR 28-APR-1999; 99US-0131449P.

PR 30-APR-1999; 99US-0132048P.

PR 30-APR-1999; 99US-0132407P.

PR 04-MAY-1999; 99US-0132484P.

PR 05-MAY-1999; 99US-0132485P.

PR 06-MAY-1999; 99US-0132486P.

PR 07-MAY-1999; 99US-0132863P.

PR 11-MAY-1999; 99US-0134256P.

PR 14-MAY-1999; 99US-0134218P.

PR 14-MAY-1999; 99US-0134219P.

PR 14-MAY-1999; 99US-0134221P.

PR 14-MAY-1999; 99US-0134370P.

PR 18-MAY-1999; 99US-0134768P.

PR 19-MAY-1999; 99US-0134941P.

PR 20-MAY-1999; 99US-0135124P.

PR 21-MAY-1999; 99US-0135353P.

PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 52.4%; Score 43; DB 3; Length 240;
Best Local Similarity 61.5%; Pred. No. 49;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 3 PFPKLVVEVFPFP 15
| | | | |
Db 203 PTPHLVVEITFPFP 215

Search completed: April 12, 2005, 08:18:27
Job time : 86 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 12, 2005, 07:49:30 ; Search time 17.5909 Seconds
(without alignments)
63.654 Million cell updates/sec

Title: US-10-691-330-8

Perfect score: 82

Sequence: 1 LKPPFKLKVEVFPFP 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	82	100.0	15	4	US-09-641-803-8
2	41	50.0	176	4	US-09-248-796A-16790
3	41	50.0	495	4	US-09-248-796A-19503
4	40	48.8	96	4	US-09-583-110-4463
5	40	48.8	451	4	US-09-555-820A-16
6	40	48.8	462	4	US-09-555-820A-20
7	40	48.8	618	4	US-09-555-820A-14
8	40	48.8	626	4	US-09-214-619-4
9	40	48.8	626	4	US-09-555-820A-12
10	39	47.6	262	4	US-09-248-796A-16764
11	39	47.6	346	4	US-09-328-352-5541
12	39	47.6	642	4	US-09-328-352-5199
13	39	47.6	759	2	US-08-637-759B-89
14	39	47.6	759	3	US-08-871-355A-89
15	39	47.6	759	3	US-09-201-945-89
16	38	46.3	164	4	US-09-621-976-5565
17	38	46.3	178	4	US-09-328-352-7054
18	38	46.3	312	3	US-08-793-331-2
19	38	46.3	595	4	US-09-949-016-9801
20	38	46.3	724	3	US-08-793-331-4
21	38	46.3	725	3	US-08-793-331-6
22	38	46.3	1536	3	US-09-413-814-10
23	37	45.1	69	4	US-09-248-796A-21731
24	37	45.1	71	4	US-09-328-352-5884
25	37	45.1	142	4	US-09-248-796A-15390
26	37	45.1	162	5	PCT-US93-03076-8
27	37	45.1	191	4	US-09-134-000C-6524

28 37 45.1 231 4 US-09-107-532A-5301 Sequence 5301, Ap
29 37 45.1 358 4 US-09-270-767-44738 Sequence 44738, A
30 37 45.1 422 4 US-09-248-796A-14166 Sequence 14166, A
31 37 45.1 485 4 US-09-252-991A-30297 Sequence 30297, A
32 37 45.1 512 4 US-09-270-767-42979 Sequence 42979, A
33 37 45.1 517 4 US-09-252-991A-21294 Sequence 21294, A
34 37 45.1 706 4 US-09-248-796A-19007 Sequence 19007, A
35 37 45.1 769 4 US-09-949-016-10665 Sequence 10665, A
36 37 45.1 780 4 US-09-949-016-6372 Sequence 6372, Ap
37 37 45.1 861 4 US-09-826-312A-10 Sequence 10, Appl
38 37 45.1 861 4 US-09-542-497A-10 Sequence 10, Appl
39 37 45.1 943 2 US-08-808-982-7 Sequence 7, Appl
40 37 45.1 943 3 US-09-306-902A-7 Sequence 7, Appl
41 37 45.1 1253 3 US-08-864-785-2 Sequence 2, Appl
42 37 45.1 1513 5 PCT-US93-03076-2 Sequence 2, Appl
43 37 45.1 2756 1 US-08-375-709-11 Sequence 11, Appl
44 37 45.1 2756 1 US-08-752-929-11 Sequence 11, Appl
45 37 45.1 2756 3 US-09-090-793-7 Sequence 7, Appl

ALIGNMENTS

RESULT 1

US-09-641-803-8

; Sequence 8, Application US/09641803

; Patent No. 6500798

; GENERAL INFORMATION:

; APPLICANT: STANTON, G. John

; APPLICANT: HUGHES, Thomas K.

; APPLICANT: BOLDOGH, Istvan

; TITLE OF INVENTION: USE OF COLOSTRININ, CONSTITUENT PEPTIDES THEREOF, AND ANALOGS THEREOF AS OXIDATIVE STRESS REGULATORS

; TITLE OF INVENTION: ANALOGS THEREOF AS OXIDATIVE STRESS REGULATORS

; FILE REFERENCE: 265.00220101

; CURRENT APPLICATION NUMBER: US/09/641.803

; CURRENT FILING DATE: 2000-08-17

; PRIOR FILING DATE: 1999-08-17

; NUMBER OF SEQ ID NOS: 34

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 8

; LENGTH: 15

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE: Description of Artificial Sequence: synthetic

; OTHER INFORMATION: peptide

; OTHER INFORMATION: peptide

US-09-641-803-8

Query Match 100.0%; Score 82; DB 4; Length 15;

Best Local Similarity 100.0%; Pred. No. 6.5e-07;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LKPPFKLKVEVFPFP 15

Db 1 LKPPFKLKVEVFPFP 15

RESULT 2

US-09-248-796A-16790

; Sequence 16790, Application US/09248796A

; Patent No. 6747137

; GENERAL INFORMATION:

; APPLICANT: Keith Weinstock et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN

; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.132

; CURRENT APPLICATION NUMBER: US/09/248.796A

; CURRENT FILING DATE: 1999-02-12

; PRIOR APPLICATION NUMBER: US 60/074,725

; PRIOR FILING DATE: 1998-02-13

; PRIOR APPLICATION NUMBER: US 60/096,409

; PRIOR FILING DATE: 1998-08-13

us-10-691-330-8-rai

Tue Apr 12 09:05:00 2005

```

; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 16790
; LENGTH: 176
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-16790

Query Match      50.0%; Score 41; DB 4; Length 176;
Best Local Similarity 53.8%; Pred. No. 30;
Matches 7; Conservative 3; Mismatches 0; Gaps 0;

Qy 1 LKPPFKLKVVEFP 13
Db 137 LKDPFKAEVKILP 149

RESULT 3
US-09-248-796A-19503
; Sequence 19503, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 19503
; LENGTH: 495
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-19503

Query Match      50.0%; Score 41; DB 4; Length 495;
Best Local Similarity 66.7%; Pred. No. 88;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 4 FPKLKVEVPFP 15
Db 14 FYTLKVKCFPP 25

RESULT 4
US-09-583-110-4463
; Sequence 4463, Application US/09583110
; Patent No. 6699703
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; TITLE OF INVENTION: Pneumoniae for Diagnostics and Therapeutics
; FILE REFERENCE: PATH00-07A
; CURRENT APPLICATION NUMBER: US/09/583,110
; CURRENT FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 4463
; LENGTH: 96
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-4463

Query Match      48.8%; Score 40; DB 4; Length 96;
Best Local Similarity 50.0%; Pred. No. 24;
Matches 7; Conservative 3; Mismatches 0; Gaps 0;

Qy 4 FPKLKVEVPFP 15
Db 14 FYTLKVKCFPP 25

us-10-691-330-8-rai

Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 2 KPPFKLKVVEVPFP 15
Db 7 EPLIKFKIVVFPFP 20

RESULT 5
US-09-555-820A-16
; Sequence 16, Application US/09555820A
; Patent No. 6680429
; GENERAL INFORMATION:
; APPLICANT: Webster, David
; APPLICANT: Burgess, Diane
; TITLE OF INVENTION: A Starchless Variety of Pisum Sativum Having Elevated Levels of St
; FILE REFERENCE: SVS3801P0302US
; CURRENT APPLICATION NUMBER: US/09/555,820A
; CURRENT FILING DATE: 2000-08-29
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 16
; LENGTH: 451
; TYPE: PRT
; ORGANISM: Pisum sativum
US-09-555-820A-16

Query Match      48.8%; Score 40; DB 4; Length 451;
Best Local Similarity 63.6%; Pred. No. 1.2e+02;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 4 FPKLKVEVPFP 14
Db 31 FPSFKVQNFPP 41

RESULT 6
US-09-555-820A-20
; Sequence 20, Application US/09555820A
; Patent No. 6680429
; GENERAL INFORMATION:
; APPLICANT: Webster, David
; APPLICANT: Burgess, Diane
; TITLE OF INVENTION: A Starchless Variety of Pisum Sativum Having Elevated Levels of St
; FILE REFERENCE: SVS3801P0302US
; CURRENT APPLICATION NUMBER: US/09/555,820A
; CURRENT FILING DATE: 2000-08-29
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20
; LENGTH: 462
; TYPE: PRT
; ORGANISM: Pisum sativum
US-09-555-820A-20

Query Match      48.8%; Score 40; DB 4; Length 462;
Best Local Similarity 63.6%; Pred. No. 1.2e+02;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 4 FPKLKVEVPFP 14
Db 33 FPSFKVQNFPP 43

RESULT 7
US-09-555-820A-14
; Sequence 14, Application US/09555820A
; Patent No. 6680429
; GENERAL INFORMATION:
; APPLICANT: Webster, David
; APPLICANT: Burgess, Diane
; TITLE OF INVENTION: A Starchless Variety of Pisum Sativum Having Elevated Levels of St
; FILE REFERENCE: SVS3801P0302US
; CURRENT APPLICATION NUMBER: US/09/555,820A
```

; CURRENT FILING DATE: 2000-08-29
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; LENGTH: 618
; TYPE: PRT
; ORGANISM: Pisum sativum
US-09-555-820A-14

Query Match 48.8%; Score 40; DB 4; Length 618;
Best Local Similarity 63.6%; Pred. No. 1.6e+02;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 4 FPKLKVEVFPF 14
|||:||||
Db 33 FPFKQVNFPP 43

RESULT 8

US-09-214-619-4
; Sequence 4, Application US/09214619
; Patent No. 6538180
; GENERAL INFORMATION:

; APPLICANT:
; TITLE OF INVENTION: METHOD FOR INCREASING SUCROSE
; TITLE OF INVENTION: CONTENT OF PLANTS
; NUMBER OF SEQUENCES: 4
; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/214,619

; FILING DATE:

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 626 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-09-214-619-4

Query Match 48.8%; Score 40; DB 4; Length 626;
Best Local Similarity 63.6%; Pred. No. 1.6e+02;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 4 FPKLKVEVFPF 14
|||:||||
Db 33 FPFKQVNFPP 43

RESULT 9

US-09-555-820A-12
; Sequence 12, Application US/09555820A
; Patent No. 6680429
; GENERAL INFORMATION:

; APPLICANT: Webster, David

; TITLE OF INVENTION: A Starchless Variety of Pisum Sativum Having Elevated Levels of S

; FILE REFERENCE: SVS3801P0302US

; CURRENT APPLICATION NUMBER: US/09/555,820A

; CURRENT FILING DATE: 2000-08-29

; NUMBER OF SEQ ID NOS: 21

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 12

; LENGTH: 626

; TYPE: PRT

; ORGANISM: Pisum sativum

US-09-555-820A-12

Query Match 48.8%; Score 40; DB 4; Length 626;
Best Local Similarity 63.6%; Pred. No. 1.6e+02;

Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
Qy 4 FPKLKVEVFPF 14
|||:||||
Db 33 FPFKQVNFPP 43

RESULT 10

US-09-248-796A-16764
; Sequence 16764, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:

; APPLICANT: Keith Weinstock et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN

; FILE REFERENCE: 107196.132

; CURRENT APPLICATION NUMBER: US/09/248,796A

; CURRENT FILING DATE: 1999-02-12

; PRIOR APPLICATION NUMBER: US 60/074,725

; PRIOR FILING DATE: 1998-02-13

; PRIOR APPLICATION NUMBER: US 60/096,409

; PRIOR FILING DATE: 1998-08-13

; NUMBER OF SEQ ID NOS: 28208

; SEQ ID NO 16764

; LENGTH: 262

; TYPE: PRT

; ORGANISM: Candida albicans

US-09-248-796A-16764

Query Match 47.6%; Score 39; DB 4; Length 262;
Best Local Similarity 54.5%; Pred. No. 95;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 2 KPPPKLKVEVPF 12
|:||||:
Db 65 KEFPDLKIEY 75

RESULT 11

US-09-328-352-5541
; Sequence 5541, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:

; APPLICANT: Gary L. Breton et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER

; FILE REFERENCE: GFC99-03PA

; CURRENT APPLICATION NUMBER: US/09/328,352

; CURRENT FILING DATE: 1999-06-04

; NUMBER OF SEQ ID NOS: 8252

; SEQ ID NO 5541

; LENGTH: 346

; TYPE: PRT

; ORGANISM: Acinetobacter baumannii

US-09-328-352-5541

Query Match 47.6%; Score 39; DB 4; Length 346;
Best Local Similarity 88.9%; Pred. No. 1.3e+02;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LKPPPKLKV 9
|||:||||
Db 164 LKQFPKLKV 172

RESULT 12

US-09-328-352-5199
; Sequence 5199, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:

; APPLICANT: Gary L. Breton et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER

; FILE REFERENCE: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS

us-10-691-330-8.ra1

Tue Apr 12 09:05:00 2005

```
FILE REFERENCE: GTC99-03PA
CURRENT APPLICATION NUMBER: US/09/328,352
CURRENT FILING DATE: 1999-06-04
NUMBER OF SEQ ID NOS: 8252
SEQ ID NO 5199
LENGTH: 642
TYPE: PRT
ORGANISM: Acinetobacter baumannii
US-09-328-352-5199

Query Match 47.6%; Score 39; DB 4; Length 642;
Best Local Similarity 46.7%; Pred. No. 2.4e+02;
Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 LKPPKLVKVEVFPF 15
Db 140 LKPLSKLIEQYPLP 154

RESULT 13
US-08-637-759B-89
Sequence 89, Application US/08637759B
Patent No. 5876931
GENERAL INFORMATION:
APPLICANT: David William Holden
TITLE OF INVENTION: Identification of Genes
NUMBER OF SEQUENCES: 501
CORRESPONDENCE ADDRESS:
ADDRESSEE: Patrea L. Pabst
STREET: 2800 One Atlantic Center
STREET: 1201 West Peachtree Street
CITY: Atlanta
STATE: Georgia
COUNTRY: USA
ZIP: 30309-3450
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/637,759B
FILING DATE: 03-MAY-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB95/02875
FILING DATE: 11-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: RPMS 101 CON
TELEPHONE: (404) 873-8794
TELEFAX: (404) 873-8795
INFORMATION FOR SEQ ID NO: 89:
SEQUENCE CHARACTERISTICS:
LENGTH: 759 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
US-08-637-759B-89

Query Match 47.6%; Score 39; DB 2; Length 759;
Best Local Similarity 46.2%; Pred. No. 2.8e+02;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 3 PFPKLVKVEVFPF 15
Db 462 PLPEVNIEVLPEP 474

RESULT 15
US-09-201-945-89
Sequence 89, Application US/09201945
Patent No. 6342215
GENERAL INFORMATION:
APPLICANT: David William Holden
TITLE OF INVENTION: Identification of Genes
NUMBER OF SEQUENCES: 501
CORRESPONDENCE ADDRESS:
ADDRESSEE: Patrea L. Pabst
STREET: 2800 One Atlantic Center
STREET: 1201 West Peachtree Street
CITY: Atlanta
STATE: Georgia
COUNTRY: USA
ZIP: 30309-3450
COMPUTER READABLE FORM:
```

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/201,945
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/537,759
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: RPLS 101
TELEPHONE: (404) 873-8794
TELEFAX: (404) 873-8795
INFORMATION FOR SEQ ID NO: 89:
SEQUENCE CHARACTERISTICS:
LENGTH: 759 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
US-09-201-945-89

Query Match 47.6%; Score 39; DB 3; Length 759;
Best Local Similarity 46.2%; Pred. NO. 2.8e+02;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Oy 3 PFPKLKVEVPFP 15
Db 462 PLPEVNIEVLPEP 474

Search completed: April 12, 2005, 08:07:54
Job time : 18.5909 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 12, 2005, 07:49:30 ; Search time 46.5 Seconds
(without alignments)
107.096 Million cell updates/sec.

Title: US-10-691-330-8

Perfect score: 82

Sequence: 1 LKPPPKLVKVEVFPP 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1418010 seqs, 33197259 residues

Total number of hits satisfying chosen parameters: 1418010

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	82	100.0	15	14 US-10-281-652-8	Sequence 8, Appli
2	82	100.0	15	17 US-10-691-330-8	Sequence 8, Appli
3	48.5	59.1	309	15 US-10-424-599-23568	Sequence 235688
4	46	56.1	100	15 US-10-424-599-146821	Sequence 146821
5	46	56.1	365	16 US-10-437-963-178880	Sequence 178880
6	44.5	54.3	1173	16 US-10-437-963-172791	Sequence 172791
7	43	52.4	954	16 US-10-771-911-3	Sequence 3, Appli
8	42.5	51.8	1515	16 US-10-437-963-137588	Sequence 137588
9	42	51.2	85	15 US-10-424-599-266211	Sequence 266211
10	42	51.2	232	15 US-10-104-047-3111	Sequence 3111, Ap
11	42	51.2	300	16 US-10-437-963-117046	Sequence 117046
12	42	51.2	410	15 US-10-425-114-64363	Sequence 64363, A
13	42	51.2	443	15 US-10-425-114-45491	Sequence 45491, A

14	42	51.2	554	15	US-10-108-260A-2682	Sequence 2682, Ap
15	42	51.2	950	15	US-10-295-027-1340	Sequence 1340, Ap
16	42	51.2	1034	16	US-10-437-963-201195	Sequence 201195
17	41.5	50.6	1675	16	US-10-826-909-34	Sequence 34, Appl
18	41	50.0	59	15	US-10-424-599-199084	Sequence 199084
19	41	50.0	68	15	US-10-424-599-224398	Sequence 224398
20	41	50.0	74	9	US-09-738-626-6030	Sequence 6030, Ap
21	41	50.0	90	15	US-10-424-599-274254	Sequence 274254
22	41	50.0	188	10	US-09-791-932-90	Sequence 90, Appl
23	41	50.0	239	15	US-10-425-114-66814	Sequence 66814, A
24	41	50.0	267	16	US-10-767-701-44378	Sequence 44378, A
25	41	50.0	333	15	US-10-424-599-213540	Sequence 213540
26	41	50.0	374	16	US-10-437-963-118336	Sequence 118336
27	41	50.0	764	16	US-10-437-963-170833	Sequence 170833
28	41	50.0	842	16	US-10-437-963-193321	Sequence 193321
29	40.5	49.4	1556	15	US-10-369-493-1796	Sequence 1796, Ap
30	40	48.8	105	16	US-10-437-963-144851	Sequence 144851
31	40	48.8	106	15	US-10-424-599-150187	Sequence 150187
32	40	48.8	121	15	US-10-424-599-218369	Sequence 218369
33	40	48.8	121	16	US-10-767-701-59034	Sequence 59034, A
34	40	48.8	131	16	US-10-437-963-138858	Sequence 138858
35	40	48.8	171	15	US-10-425-114-48415	Sequence 48415, A
36	40	48.8	201	16	US-10-767-701-46892	Sequence 46892, A
37	40	48.8	258	15	US-10-424-599-176441	Sequence 176441
38	40	48.8	304	15	US-10-424-599-170626	Sequence 170626
39	40	48.8	340	14	US-10-083-357-1337	Sequence 1337, Ap
40	40	48.8	496	15	US-10-425-114-40273	Sequence 40273, A
41	40	48.8	626	9	US-09-906-209-12	Sequence 12, Appl
42	40	48.8	626	9	US-09-906-209-13	Sequence 13, Appl
43	40	48.8	650	16	US-10-437-963-187658	Sequence 187658
44	40	48.8	711	15	US-10-424-599-280508	Sequence 280508
45	40	48.8	927	16	US-10-437-963-121736	Sequence 121736

ALIGNMENTS

RESULT 1

US-10-281-652-8
; Sequence 8, Application US/10281652
; Publication No. US20030091606A1
; GENERAL INFORMATION:
; APPLICANT: STANTON, G. John
; APPLICANT: HUGHES, Thomas K.
; APPLICANT: BOLDOGH, Istvan
; TITLE OF INVENTION: USE OF COLOSTRININ, CONSTITUENT PEPTIDES THEREOF, AND
; FILE REFERENCE: 265.00220101
; CURRENT APPLICATION NUMBER: US/10/281.652
; CURRENT FILING DATE: 2002-10-28
; PRIOR APPLICATION NUMBER: US/09/641.803
; PRIOR FILING DATE: 2000-08-17
; PRIOR APPLICATION NUMBER: 60/149,310
; PRIOR FILING DATE: 1999-08-17
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: peptide
US-10-281-652-8

Query Match 100.0%; Score 82; DB 14; Length 15;

Best Local Similarity 100.0%; Pred. No. 2.6e-06;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LKPPPKLVKVEVFPP 15

Db 1 LKPPPKLVKVEVFPP 15

RESULT 2
US-10-691-330-8
; Sequence 8, Application US/10691330
; Publication No. US20050042300A1
; GENERAL INFORMATION:
; APPLICANT: Boldogh, Istvan
; APPLICANT: Stanton, G. John
; APPLICANT: Georgiades, Jerzy A.
; APPLICANT: Hughes, Thomas K., Jr.
; APPLICANT: Krugel, Marian
; TITLE OF INVENTION: USE OF COLOSTRININ, CONSTITUENT PEPTIDES THEREOF, AND ANALOGS
; TITLE OF INVENTION: THEREOF AS INHIBITORS OF APOPTOSIS AND OTHER CELLULAR DAMAGE
; FILE REFERENCE: 265.00390101
; CURRENT APPLICATION NUMBER: US/10/691,330
; CURRENT FILING DATE: 2003-10-22
; PRIOR APPLICATION NUMBER: US 60/420,369
; PRIOR FILING DATE: 2002-10-22
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 8
; LENGTH: 15
; TYPE: PRT
; ORGANISM: artificial
; FEATURE:
; OTHER INFORMATION: Synthetic Peptides
US-10-691-330-8

Query Match 100.0%; Score 82; DB 17; Length 15;
Best Local Similarity 100.0%; Pred. No. 2.6e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LKPPFKLKVEVFPFP 15
| | | | | | | | | | | | | | |
Db 1 LKPPFKLKVEVFPFP 15

RESULT 3
US-10-424-599-235668
; Sequence 235668, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 235668
; LENGTH: 309
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(309)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_54835C.1.pep
US-10-424-599-235668

Query Match 59.1%; Score 48.5; DB 15; Length 309;
Best Local Similarity 50.0%; Pred. No. 13;
Matches 12; Conservative 2; Mismatches 1; Indels 9; Gaps 2;

Qy 1 LKPF-PKLKVEVFP-----PFP 15
| | | | | | | | | | | | | | |
Db 275 LKPFAPKPIEVFLFAIRKTLFPF 298

RESULT 4
US-10-424-599-146821
; Sequence 146821, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 146821
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_103599C.1.pep
US-10-424-599-146821

Query Match 56.1%; Score 46; DB 15; Length 100;
Best Local Similarity 72.7%; Pred. No. 9.9;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 3 PPKLKVEVFP 13
| | | | | | | | | | | | |
Db 79 PPKIKVKVSP 89

RESULT 5
US-10-437-963-178880
; Sequence 178880, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Brad
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 178880
; LENGTH: 365
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_76396C.1.pep
US-10-437-963-178880

Query Match 56.1%; Score 46; DB 16; Length 365;
Best Local Similarity 53.3%; Pred. No. 39;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps --0;

Qy 1 LKPPFKLKVEVFPFP 15
| | | | | | | | | | | | | | |
Db 19 LLFPFKVSVQVYIVP 33

RESULT 6
US-10-437-963-172791
; Sequence 172791, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 172791
; LENGTH: 1173
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)...(1173)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_70893C.1.pep
US-10-437-963-172791

Query Match 54.3%; Score 44.5; DB 16; Length 1173;
Best Local Similarity 60.0%; Pred. No. 2.3e+02;
Matches 9; Conservative 2; Mismatches 3; Indels 1; Gaps 1;

Qy 2 KPFPKLKVEVF-PFP 15
Db 968 RPFPLAKFYFGPFP 982

RESULT 7
US-10-771-931-3
; Sequence 3, Application US/10771931
; Publication No. US20040151737A1
; GENERAL INFORMATION:
; APPLICANT: Courtney, Harry
; TITLE OF INVENTION: Streptococcal Serum Opacity Factors And Fibronectin-Binding Prote
; FILE REFERENCE: 13314.1001U
; CURRENT APPLICATION NUMBER: US/10/771,931
; CURRENT FILING DATE: 2004-02-04
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 954
; TYPE: PRT
; ORGANISM: Streptococcus pyogenes
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(954)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-10-771-931-3

Query Match 52.4%; Score 43; DB 16; Length 954;
Best Local Similarity 53.8%; Pred. No. 3.2e+02;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 3 PFPKLKVEVFPP 15
Db 730 PIPELDIEWVPI 742

RESULT 8
US-10-437-963-137588
; Sequence 137588, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 137588
; LENGTH: 1515
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)...(1515)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_39057C.1.pep
US-10-437-963-137588

Query Match 51.8%; Score 42.5; DB 16; Length 1515;
Best Local Similarity 60.0%; Pred. No. 6.3e+02;
Matches 9; Conservative 2; Mismatches 3; Indels 1; Gaps 1;

Qy 2 KPFPKLKVEVF-PFP 15
Db 1374 QPFPKLKVEVFPGFP 1388

RESULT 9
US-10-424-599-266211
; Sequence 266211, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 266211
; LENGTH: 85
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_8240C.1.pep
US-10-424-599-266211

Query Match 51.2%; Score 42; DB 15; Length 85;
Best Local Similarity 58.3%; Pred. No. 36;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 LKPPPKLKVEVF 12
Db 13 ISPFKLQVEIF 24

RESULT 10
US-10-104-047-3111
; Sequence 3111, Application US/10104047
; Publication No. US20030236392A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20030236392A1e1 full length cDNA

; FILE REFERENCE: HI-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3111
; LENGTH: 232
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-104-047-3111

Query Match 51.2%; Score 42; DB 15; Length 232;
Best Local Similarity 53.3%; Pred. No. 1e+02; 5; Indels 0; Gaps 0;
Matches 8; Conservative 2; Mismatches 5;

Qy 1 LKFPFKLKVEVFPF 15
| : | | | : | | |
Db 207 LRKFFVLVHPVFPF 221

RESULT 11
US-10-437-963-117046
; Sequence 117046, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 117046
; LENGTH: 300
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_2048C.1.pep
US-10-437-963-117046

Query Match 51.2%; Score 42; DB 16; Length 300;
Best Local Similarity 54.5%; Pred. No. 1.4e+02; 2; Indels 0; Gaps 0;
Matches 6; Conservative 3; Mismatches 2;

Qy 4 FPKLKVEVFPF 14
| : | | | : | | |
Db 46 FPSLRFEIYFPF 56

RESULT 12
US-10-425-114-64363
; Sequence 64363, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114

; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 64363
; LENGTH: 410
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3689-218-D2_FLI.pep
US-10-425-114-64363

Query Match 51.2%; Score 42; DB 15; Length 410;
Best Local Similarity 54.5%; Pred. No. 1.9e+02; 3; Indels 0; Gaps 0;
Matches 6; Conservative 3; Mismatches 2;

Qy 4 FPKLKVEVFPF 14
| : | | | : | | |
Db 159 FPSLRFEIYFPF 169

RESULT 13
US-10-425-114-45491
; Sequence 45491, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 45491
; LENGTH: 443
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700377865_FLI.pep
US-10-425-114-45491

Query Match 51.2%; Score 42; DB 15; Length 443;
Best Local Similarity 54.5%; Pred. No. 2.1e+02; 3; Indels 0; Gaps 0;
Matches 6; Conservative 3; Mismatches 2;

Qy 4 FPKLKVEVFPF 14
| : | | | : | | |
Db 193 FPSLRFEIYFPF 203

RESULT 14
US-10-108-260A-2682
; Sequence 2682, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: NO. US20040005560A1el full length cDNA
; FILE REFERENCE: HI-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2682
; LENGTH: 554
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-108-260A-2682

Query Match 51.2%; Score 42; DB 15; Length 554;
Best Local Similarity 66.7%; Pred. No. 2.6e+02;

Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 LKPPPKLKVEVF 12
| | | | |
Db 83 LDPLPSLKVEVY 94

RESULT 15

US-10-295-027-1340
; Sequence 1340, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1340
; LENGTH: 950
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-1340

Query Match 51.2%; Score 42; DB 15; Length 950;
Best Local Similarity 66.7%; Pred. No. 4.6e+02;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 LKPPPKLKVEVF 12
| | | | |
Db 479 LDPLPSLKVEVY 490

Search completed: April 12, 2005, 07:55:22
Job time : 46.5 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 12, 2005, 07:49:31 ; Search time 10.7727 Seconds
(without alignments)
133.973 Million cell updates/sec

Title: US-10-691-330-8

Perfect score: 82

Sequence: 1 LKPPFKLKVEVFPFP 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_79:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	43	52.4	127	F86175	protein F19P19.17
2	43	52.4	790	G90477	hypothetical prote
3	43	52.4	1007	C84668	probable receptor-
4	42	51.2	190	T45013	hypothetical prote
5	42	51.2	343	T28908	hypothetical prote
6	42	51.2	428	C97432	probable oxidoredu
7	42	51.2	428	A02650	oxidoreductase ord
8	42	51.2	939	T18974	hypothetical prote
9	42	51.2	965	T22933	hypothetical prote
10	41.5	50.6	379	S44745	CO2D5.3 protein -
11	41.5	50.6	1675	LRRTH	clathrin heavy cha
12	41	50.0	398	T38233	probable cystathio
13	41	50.0	401	A82221	extracellular solu
14	41	50.0	496	S75790	hypothetical prote
15	41	50.0	537	C64432	hypothetical prote
16	41	50.0	736	D86846	competence protein
17	41	50.0	1353	T19157	probable metal bin
18	41	50.0	1424	T19156	probable metal bin
19	40.5	49.4	1556	S39393	probable membrane
20	40	48.8	134	T32913	hypothetical prote
21	40	48.8	134	C87723	protein C54G6.4 [i
22	40	48.8	193	E86233	hypothetical prote
23	40	48.8	212	A82992	hypothetical prote
24	40	48.8	212	H98291	hypothetical prote
25	40	48.8	259	1 YX517	thymidylate syntha
26	40	48.8	336	B75188	probable 4-hydroxy
27	40	48.8	424	T33839	hypothetical prote
28	40	48.8	572	T34345	hypothetical prote
29	39	47.6	258	T51898	hypothetical prote

30	39	47.6	307	1 YXMST	thymidylate syntha
31	39	47.6	327	2 JS0402	gamma-gliadin prec
32	39	47.6	371	2 D97042	hypothetical prote
33	39	47.6	385	2 G90274	hypothetical prote
34	39	47.6	481	2 C64945	probable rRNA meth
35	39	47.6	481	2 A98947	probable rRNA meth
36	39	47.6	481	2 E85795	probable nucleolar
37	39	47.6	509	2 T50006	receptor protein k
38	39	47.6	562	2 T05758	hypothetical prote
39	39	47.6	681	2 AF0697	probable type III
40	39	47.6	706	2 G71004	hypothetical prote
41	39	47.6	782	2 A82940	hypothetical prote
42	39	47.6	1263	2 AH2011	heterocyst glycoli
43	39	47.6	1791	2 T02345	hypothetical prote
44	38.5	47.0	270	2 AD1408	hypothetical prote
45	38.5	47.0	270	2 AD1784	hypothetical prote

ALIGNMENTS

RESULT 1

F86175
protein F19P19.17 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C:Accession: F86175
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: F86175
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-127 <STO>
A:Cross-references: UNIPROT:O22688; GB:AE005172; NID:g2341037; PIDN:AAB70437.1; GSPDB:GN
C:Genetics:
A:Gene: F19P19.17
A:Map position: 1

Query Match 52.4%; Score 43; DB 2; Length 127;
Best Local Similarity 80.0%; Pred. No. 4.8;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LKPPFKLKVE 10
|||||:
Db 59 LKPPFKLKSE 68

RESULT 2

G90477
hypothetical protein soxM [imported] - Sulfolobus solfataricus
C:Species: Sulfolobus solfataricus
C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 09-Jul-2004
C:Accession: G90477
R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, P.
arrett, R.A.; Ragan, M.A.; Senses, C.W.; Van der Oost, J.
submitted to GenBank, April 2001
A:Description: Sulfolobus solfataricus complete genome.
A:Reference number: A99139
A:Accession: G90477
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-790 <KUR>
A:Cross-references: UNIPROT:Q97UN0; GB:AE006641; NID:gl3816356; PIDN:AAK43078.1; GSPDB:G

C;Genetics:
A;Gene: soxM
C;Superfamily: cytochrome-c oxidase chain I/III; cytochrome-c oxidase chain I homology
C;Keywords: copper; electron transfer; heme; iron; magnesium; membrane-associated complex
F;63,377/Binding site: heme a iron (His) (axial ligands) #status predicted
F;239,289,290/Binding site: copper (His) #status predicted
F;239-243/Cross-link: 1'-histidyl-3'-tyrosine (His-Tyr) #status predicted
F;243/Binding site: oxygen (Tyr) #status predicted
F;375/Binding site: heme a3 iron (His) (axial ligand) #status predicted

Query Match 52.4%; Score 43; DB 2; Length 790;
Best Local Similarity 72.7%; Pred. No. 31; Indels 0; Gaps 0;
Matches 8; Conservative 0; Mismatches 3

Qy 5 PKLKVEVFPFP 15
Db 650 PPLKVEVFPFP 660

RESULT 3
C84668
probable receptor-like protein kinase [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C;Accession: C84668
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.;
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; UID:20083487; PMID:10617197
A;Accession: C84668
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1007 <STO>
A;Cross-references: UNIPROT:Q9ZVD4; GB:AE002093; NID:G3885336; PIDN:AAC77864.1; GSPDB:GN
C;Genetics:
A;Gene: At2g27060
A;Map position: 2

Query Match 52.4%; Score 43; DB 2; Length 1007;
Best Local Similarity 53.8%; Pred. No. 40; Indels 0; Gaps 0;
Matches 7; Conservative 3; Mismatches 3

Qy 2 KPFPKLKVEVFPFP 14
Db 898 KPFPKLKVEVFPFP 910

RESULT 4
T45013
hypothetical protein [imported] - Methanosarcina acetivorans plasmid pc2A
C;Species: Methanosarcina acetivorans
C;Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 09-Jul-2004
C;Accession: T45013
R;Metcalf, W.W.; Zhang, J.K.; Apollinario, E.; Sowers, K.R.; Wolfe, R.S.
Proc. Natl. Acad. Sci. U.S.A. 94, 2626-2631, 1997
A;Title: A genetic system for Archaea of the genus Methanosarcina: liposome-mediated tra
A;Reference number: Z22897; UID:97226004; PMID:9122246
A;Accession: T45013
A;Status: preliminary; translated from GB/EMBL/DBDJ
A;Molecule type: DNA
A;Residues: 1-190 <MET>
A;Cross-references: UNIPROT:P94913; EMBL:U78295; NID:G1763609; PIDN:AAB39747.1; PID:G176
C;Genetics:
A;Genome: plasmid pc2A

Query Match 51.2%; Score 42; DB 2; Length 190;
Best Local Similarity 64.3%; Pred. No. 11; Indels 1; Gaps 1;
Matches 9; Conservative 2; Mismatches 1

Qy 4 FPKLKVEVFPFP--FP 15

Db 110 FPKLKELFPPEQFP 123

RESULT 5
T28908
hypothetical protein T26C11.2 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C;Accession: T28908
R;Martin, J.
submitted to the EMBL Data Library, December 1995
A;Description: The sequence of C. elegans cosmid T26C11.
A;Reference number: Z20542
A;Accession: T28908
A;Status: preliminary; translated from GB/EMBL/DBDJ
A;Molecule type: DNA
A;Residues: 1-343 <MAR>
A;Cross-references: UNIPROT:Q22807; EMBL:U41017; PIDN:AAC48211.1; GSPDB:GN00028; CESP:T2;
A;Experimental source: strain Bristol N2; clone T26C11
C;Genetics:
A;Gene: CESP:T26C11.2
A;Map position: X
C;Superfamily: hydroxyproline-rich glycoprotein

Query Match 51.2%; Score 42; DB 2; Length 343;
Best Local Similarity 64.3%; Pred. No. 20; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 5

Qy 2 KPFPKLKVEVFPFP 15
Db 7 KPFPKLKVEVFPFP 20

RESULT 6
C97432
probable oxidoreductase ordl AGR_C_1066 [imported] - Agrobacterium tumefaciens (strain C
C;Species: Agrobacterium tumefaciens
C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
C;Accession: C97432
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
A.; Liu, F.; Wiliam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A;Reference number: A97359; UID:21608551; PMID:11743194
A;Accession: C97432
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-428 <KUR>
A;Cross-references: UNIPROT:Q8UH68; GB:AE007869; PIDN:AAK86412.1; PID:G15155546; GSPDB:G
C;Genetics:
A;Gene: AGR_C_1066
A;Map position: circular chromosome
C;Superfamily: hypothetical protein HI0499

Query Match 51.2%; Score 42; DB 2; Length 428;
Best Local Similarity 61.5%; Pred. No. 25; Indels 4; Gaps 0;
Matches 8; Conservative 1; Mismatches 1

Qy 1 LKFPKLKVEVFP 13
Db 393 LKFPKLKVEVFP 405

RESULT 7
AD2650
oxidoreductase ordL [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C;Accession: AD2650
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.

Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; MUID:21608550; PMID:111743193
A:Accession: AD2650
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-428 <KUR>
A:Cross-references: UNIPROT:Q8UHS8; GB:AE008688; PIDN:RAL41618.1; PID:gl7738956; GSPDB:C
A:Experimental source: strain C58 (dupont)
C:Genetics:
A:Gene: Ordl
A:Map position: circular chromosome
C:Superfamily: hypothetical protein HI0499

Query Match 51.2%; Score 42; DB 2; Length 428;
Best Local Similarity 61.5%; Pred. No. 25;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 LKPPPKLKVEVFP 13
||| :|||
Db 393 LAPFARLKVPFP 405

RESULT 8
T18974
hypothetical protein C06A1.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T18974
R:McMurray, A.
submitted to the EMBL Data Library, June 1995
A:Reference number: Z19054
A:Accession: T18974
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-939 <WIL>
A:Cross-references: EMBL:Z49886; PIDN:CAA90054.1; GSPDB:GN00020; CESP:C06A1.4
A:Experimental source: clone C06A1
C:Genetics:
A:Gene: CESP:C06A1.4
A:Map position: 2
A:Introns: 52/3; 116/2; 146/3; 282/1; 524/2; 583/1; 639/2; 697/3; 779/3; 901/2

Query Match 51.2%; Score 42; DB 2; Length 939;
Best Local Similarity 61.5%; Pred. No. 55;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 LKPPPKLKVEVFP 13
||| :|||
Db 162 LKSLPCIKLEVPF 174

RESULT 9
T22933
hypothetical protein F58G1.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T22933
R:Smye, R.
submitted to the EMBL Data Library, November 1996
A:Reference number: Z19639
A:Accession: T22933
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-965 <WIL>
A:Cross-references: UNIPROT:O62275; EMBL:Z81556; PIDN:CAB04524.1; GSPDB:GN00020; CESP:F5
A:Experimental source: clone F58G1
C:Genetics:
A:Gene: CESP:F58G1.1
A:Map position: 2
A:Introns: 52/3; 116/2; 172/3; 308/1; 550/2; 609/1; 665/2; 723/3; 805/3; 927/2

Query Match 51.2%; Score 42; DB 2; Length 965;
Best Local Similarity 61.5%; Pred. No. 57;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 LKPPPKLKVEVFP 13
||| :|||
Db 188 LKSLPCIKLEVPF 200

RESULT 10
S44745
C02D5.3 protein - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 14-Sep-1994 #sequence_revision 12-May-1995 #text_change 09-Jul-2004
C:Accession: S44745
R:Du, Z.
submitted to the EMBL Data Library, May 1993
A:Description: Sequence of the C. elegans cosmid C02D5.
A:Reference number: S44613
A:Accession: S44745
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-379 <DUZ>
A:Cross-references: UNIPROT:P34277; EMBL:L16622; NID:G289603; PID:G289606
C:Genetics:
A:Introns: 15/1; 90/1; 174/2; 196/1; 272/3

Query Match 50.6%; Score 41.5; DB 2; Length 379;
Best Local Similarity 66.7%; Pred. No. 26;
Matches 10; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

Qy 1 LKPPPKLKVEVFP 15
||| :|||
Db 320 IKDFPP-LKVESFPGP 333

RESULT 11
LRRTH
clathrin heavy chain - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 09-Jul-2004
C:Accession: A39941
R:Kirchhausen, T.; Harrison, S.C.; Ping Chow, E.; Mattaliano, R.J.; Ramachandran, K.L.; Proc. Natl. Acad. Sci. U.S.A. 84, 8805-8809, 1987
A:Title: Clathrin heavy chain: molecular cloning and complete primary structure.
A:Reference number: A39941; MUID:88097376; PMID:3480512
A:Accession: A39941
A:Molecule type: mRNA
A:Residues: 1-1675 <KIR>
A:Cross-references: UNIPROT:P11442; GB:J03583; NID:G203301; PIDN:AAA40874.1; PID:G203302
C:Comment: Clathrin, the major protein component of coated pits and vesicles, is a three r its carboxyl end. The heavy chains are also held together by noncovalent interactions.
C:Comment: The amino end of the mature protein is blocked.
C:Superfamily: clathrin heavy chain
C:Keywords: coated pits
F:1-479/Domain: amino-terminal <TER>
F:480-523/Region: link
F:524-634/Domain: distal <DIS>
F:635-638/Region: joint #status predicted
F:639-1675/Domain: proximal <PRX>

Query Match 50.6%; Score 41.5; DB 1; Length 1675;
Best Local Similarity 64.3%; Pred. No. 1.2e+02;
Matches 9; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

Qy 2 KPFPKLKVEVFP 15
||| :|||
Db 241 QPFPKAVDFP-PP 253

RESULT 12
T38233

probable cystathionine gamma-synthase - fission yeast (*Schizosaccharomyces pombe*)

A;Residues: 1-496 <Rn>
C;Species: Schistosoma haematobium
C;date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jan-2000
C;Accession: T38233
A;Note: The nucleotide sequence was submitted to the EMBL Data Library, June 1996
C;Superfamily: Synectocystis hypothetical protein sl10827
A;Cross-references: UNIPROT:Q55425; EMBL:D64003; GB:AB001339; NID:gl001200; PIDN:BAAL052;
P;Murphy J., Harris D., Wood V., Barrell B.G., Rajandream M.A.

submitted to the EMBL Data Library, February 1998
A;Reference number: Z21780
A;Accession: T38233
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-398 <MUR>
A;Cross-references: EMBL:AL021813; PIDN:CAA15988.1; GSPDB:GN00065; SPDB:SPAC23A1.14C
A;Experimental source: strain 972n; csmid c23A1

A, Map position: 1
C: Superfamily: O-succinylhomoserine (thiol)-lyase
A: Gene: SPDB:SPAC23A1.14c
C: Geneticks:
A, Experimental source: *Salmonella typhimurium* CG5912

Query Match	50.0%	Score 41;	DB 2;	Length 398;
Best Local Similarity	57.1%;	Pred. No. 34;		
Matches	8;	Conservative	2;	Mismatches
			4;	Indels
			0;	Gaps

A; Molecule type: DNA
A; Residues: 1-496 <KAN>
A; Cross-references: UNIPROT_Q55425; ENMLD64003; GB_AB001339; NID:G1001200; PIDN:BAAL052;
A; Note: the nucleotide sequence was submitted to the ENML Data Library, June 1996
C; Superfamily: Synecocystis hypothetical protein sll0827

Query Match	50.0%	Score 41;	DB 2;	Length 496;
Best Local Similarity	63.8%	Pred. NO. 42;		
Matches	7;	Conservative	1;	Mismatches 3; Indels 0; Gaps 0;
Qy	4	FPKLKVEVPFF	14	
Db	64	FPKRPVRIFPF	74	

RESULT 15
C64432
hypothetical protein MJ1060 - Methanococcus jannaschii
C/Species: Methanococcus jannaschii
C/Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 09-Jul-2004
C/Accession: C64432

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 12, 2005, 07:49:31 ; Search time 71.1818 Seconds
(without alignments)
107.909 Million cell updates/sec

Title: US-10-691-330-8

Perfect score: 82

Sequence: 1 LKPPPKLKVVEVFPFP 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48	58.5	70	Q6DUM7	Q6dum7 fragaria an
2	47	57.3	435	Q9VVQ7	Q9vvq7 drosophila
3	47	57.3	1122	Q6YXG6	Q6yxg6 homo sapien
4	47	57.3	1580	Q6ZPH3	Q6zph3 mus musculu
5	47	57.3	2443	Q26J17	Q26j17 homo sapien
6	46	56.1	159	Q948M2	Q948m2 oryza merid
7	46	56.1	161	Q93VC4	Q93vc4 oryza rufip
8	46	56.1	161	Q93VC5	Q93vc5 oryza sativ
9	46	56.1	161	Q948M0	Q948m0 oryza glumi
10	46	56.1	161	Q948M1	Q948m1 oryza barth
11	46	56.1	161	Q948M3	Q948m3 oryza glabe
12	46	56.1	163	Q948M4	Q948m4 oryza rufip
13	46	56.1	164	Q93W90	Q93w90 oryza sativ
14	46	56.1	164	Q948L6	Q948l6 oryza rufip
15	46	56.1	164	Q948L7	Q948l7 oryza rufip
16	46	56.1	164	Q948L8	Q948l8 oryza rufip
17	46	56.1	164	Q948L9	Q948l9 oryza sativ
18	46	56.1	164	Q7GBQ8	Q7gbq8 oryza sativ
19	46	56.1	164	Q7GBQ9	Q7gbq9 oryza sativ
20	46	56.1	365	Q948L4	Q948l4 oryza sativ
21	46	56.1	365	Q948L5	Q948l5 oryza sativ
22	46	56.1	367	Q948L7	Q948l7 oryza sativ
23	45	54.9	173	Q6XM21	Q6xm21 feldmannia
24	45	54.9	348	Q6XCM2	Q6vcw5 catharanthu
25	44	53.7	621	Q7M943	Q7m943 wolinnella s
26	44	53.7	650	Q8J2K1	Q8j2k1 pichia angu
27	43	52.4	101	Q9CYU9	Q9cyu9 mus musculu
28	43	52.4	127	Q22688	Q22688 arabidopsis
29	43	52.4	213	Q22688	Q22688 streptococc
30	43	52.4	249	Q9LHC4	Q9lhc4 arabidopsis
31	43	52.4	790	Q97UN0	Q97un0 sulfolobus

32 43 52.4 954 2 Q6YBU6 Q6ybu6 streptococc
33 43 52.4 1007 2 Q9ZVD4 Q9zvd4 arabadopsis
34 43 52.4 2154 2 Q74ZT5 Q74zt5 ashbya goss
35 42 51.2 183 2 Q7N2A1 Q7n2a1 photorhabdu
36 42 51.2 190 2 P94913 P94913 methanosarc
37 42 51.2 232 2 Q8NAJ2 Q8naj2 homo sapien
38 42 51.2 289 2 Q99ZW1 Q99zw1 streptococc
39 42 51.2 289 2 Q7CF58 Q7cf58 streptococc
40 42 51.2 343 2 Q22807 Q22807 caenorhabdi
41 42 51.2 373 2 Q6ZSN7 Q6zsn7 oryza sativ
42 42 51.2 428 2 Q8UHS8 Q8uhs8 agrobacteri
43 42 51.2 535 2 Q7MME7 Q7mme7 porphyromon
44 42 51.2 762 2 Q993B3 Q993b3 simian cyto
45 42 51.2 965 2 Q62275 Q62275 caenorhabdi

ALIGNMENTS

RESULT 1

Q6DUM7 PRELIMINARY; PRT; 70 AA.
AC Q6DUM7; 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Probable mitochondrial carrier protein (Fragment).
OS Fragaria ananassa (Strawberry).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Rosales; Rosaceae; Rosoideae; Fragaria.
OX NCBI_TaxID=3747;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Receptacle;
RA Balogh A., Hencz T., Kiss E., Heszky L.E.;
RL Submitted (May-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY642688; AAT66507.1; -;
FT NON_TER
FT NON_TER 1 70
SQ SEQUENCE 70 AA; 8298 MW; DB707AS86C9EE4B9 CRC64;

Query Match 58.5%; Score 48; DB 2; Length 70;
Best Local Similarity 53.3%; Pred. No. 3.2;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 1 LKPPPKLKVVEVFPFP 15
:| ||||: |||:
Db 17 IKLSPKLKLRIFPWP 31

RESULT 2

Q9VVQ7 PRELIMINARY; PRT; 435 AA.
AC Q9VVQ7; 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE CG18231-PA.
GN Drosophila melanogaster (Fruit fly).
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.E.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazey R.G., Champe M., Pfeiffer B.D.,


```

RX PubMed=14621295;
RA Okazaki N., Kikuno R., Ohara R., Inamoto S., Koseki H., Hiraoka S.,
RA Suga Y., Nagase T., Ohara O., Koga H.;
RT "Prediction of the coding sequences of mouse homologues of KIAA gene:
RT III. the complete nucleotide sequences of 500 mouse KIAA-homologous
RT cDNAs identified by screening of terminal sequences of cDNA clones
RT randomly sampled from size-fractionated libraries.";
RL DNA Res. 10:167-180(2003).
DR EMBL; AK129453; BAC98263.1; -.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0008199; F:ferric iron binding; IEA.
DR GO; GO:0006725; P:aromatic compound metabolism; IEA.
DR InterPro; IPR000627; Dioxigenase.
FT NON TER 1
SQ SEQUENCE 1580 AA; 178755 MW; EDF5BAA8DD8CD76 CRC64;

Query Match 57.3%; Score 47; DB 2; Length 1580;
Best Local Similarity 61.5%; Pred. No. 1e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 LKPPPKLKVEVFP 13
Db 262 LTPPKLKALFP 274

RESULT 5
ID Q96J17 PRELIMINARY; PRT; 2443 AA.
AC Q96J17
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE KIAA1840 protein (Fragment).
GN Name=KIAA1840;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=21245130; PubMed=11347906;
RA Nagase T., Nakayama M., Nakajima D., Kikuno R., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XX.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro.";
RL DNA Res. 8:85-95(2001).
DR EMBL; AB058743; BAB47469.2; -.
DR InterPro; IPR001360; Glyco hydro 1.
DR PROSITE; PS00572; GLYCOSYL_HYDROL_F1_1; UNKNOWN_1.
DR PROSITE; PS00037; MYB_1; UNKNOWN_1.
FT NON TER 1
SQ SEQUENCE 2443 AA; 278805 MW; 580B24253D940D1E CRC64;

Query Match 57.3%; Score 47; DB 2; Length 2443;
Best Local Similarity 61.5%; Pred. No. 1.6e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 LKPPPKLKVEVFP 13
Db 1120 LTPPKLKALFP 1132

RESULT 6
ID Q948M2 PRELIMINARY; PRT; 159 AA.
AC Q948M2
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Transcription factor OSH3 (Fragment).
GN Name=OSH3;

```

```

OS Oryza meridionalis.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=40149;
RN [1]
RP SEQUENCE FROM N.A.
RA Sato Y., Fukuda Y., Hirano H.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB071627; BAB68273.1; -.
DR Gramene; Q948M2; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR InterPro; IPR005540; KNOX1.
DR InterPro; IPR005541; KNOX2.
DR Pfam; PF03790; KNOX1; 1.
DR Pfam; PF03791; KNOX2; 1.
FT NON TER 1
FT NON TER 159
SQ SEQUENCE 159 AA; 16897 MW; 48753DB82F95AC16 CRC64;

Query Match 56.1%; Score 46; DB 2; Length 159;
Best Local Similarity 53.3%; Pred. No. 15;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 LKPPPKLKVEVFP 15
Db 19 LLPPPKVSVQVYTP 33

RESULT 7
ID Q93VC4 PRELIMINARY; PRT; 161 AA.
AC Q93VC4
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Transcription factor OSH3 (Fragment).
GN Name=OSH3;
OS Oryza rufipogon (Wild rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4529;
RN [1]
RP SEQUENCE FROM N.A.
RA Sato Y., Fukuda Y., Hirano H.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB071658; BAB68304.1; -.
DR EMBL; AB071660; BAB68306.1; -.
DR EMBL; AB071661; BAB68307.1; -.
DR Gramene; Q93VC4; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR InterPro; IPR005540; KNOX1.
DR InterPro; IPR005541; KNOX2.
DR Pfam; PF03790; KNOX1; 1.
DR Pfam; PF03791; KNOX2; 1.
FT NON TER 1
FT NON TER 161
SQ SEQUENCE 161 AA; 17174 MW; D0C0D61C8369FA77 CRC64;

Query Match 56.1%; Score 46; DB 2; Length 161;
Best Local Similarity 53.3%; Pred. No. 15;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 LKPPPKLKVEVFP 15
Db 19 LLPPPKVSVQVYTP 33

RESULT 8
ID Q93VC5

```



```
Query Match 56.1%; Score 46; DB 2; Length 161;
Best Local Similarity 53.3%; Pred. No. 15;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 LKPPFKLKVEVFPFP 15
Db 19 LLPFPKVSQVYVTP 33

RESULT 12
Q948M4 PRELIMINARY; PRT; 163 AA.
ID Q948M4
AC Q948M4;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Transcription factor OSH3 (Fragment).
GN Name=OSH3;
OS Oryza rufipogon (Wild rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4529;
RN [1]
RP SEQUENCE FROM N.A.
RA Sato Y., Fukuda Y., Hirano H.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB071625; BAB68271.1; -.
DR Gramene; Q948M4; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR InterPro; IPR005540; KNOX1.
DR InterPro; IPR005541; KNOX2.
DR Pfam; PF03790; KNOX1; 1.
DR Pfam; PF03791; KNOX2; 1.
FT NON_TER 1
FT NON_TER 163
SQ SEQUENCE 163 AA; 17248 MW; 3E4D33040B16B866 CRC64;

Query Match 56.1%; Score 46; DB 2; Length 163;
Best Local Similarity 53.3%; Pred. No. 16;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 LKPPFKLKVEVFPFP 15
Db 18 LLPFPKVSQVYVTP 32

RESULT 13
Q93W90 PRELIMINARY; PRT; 164 AA.
ID Q93W90
AC Q93W90;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE Transcription factor OSH3 (Fragment).
GN Name=OSH3;
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RA Sato Y., Fukuda Y., Hirano H.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB071630; BAB68276.1; -.
DR EMBL; AB071635; BAB68281.1; -.
DR EMBL; AB071637; BAB68283.1; -.
DR EMBL; AB071638; BAB68284.1; -.
DR Gramene; Q93W90; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
```

```
DR InterPro; IPR005540; KNOX1.
DR InterPro; IPR005541; KNOX2.
DR Pfam; PF03790; KNOX1; 1.
DR Pfam; PF03791; KNOX2; 1.
FT NON_TER 1
FT NON_TER 164
SQ SEQUENCE 164 AA; 17403 MW; 430B9D3BAF43C6C5 CRC64;

Query Match 56.1%; Score 46; DB 2; Length 164;
Best Local Similarity 53.3%; Pred. No. 16;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 LKPPFKLKVEVFPFP 15
Db 19 LLPFPKVSQVYVTP 33

RESULT 14
Q948L6 PRELIMINARY; PRT; 164 AA.
ID Q948L6
AC Q948L6;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Transcription factor OSH3 (Fragment).
GN Name=OSH3;
OS Oryza rufipogon (Wild rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4529;
RN [1]
RP SEQUENCE FROM N.A.
RA Sato Y., Fukuda Y., Hirano H.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB071662; BAB68308.1; -.
DR Gramene; Q948L6; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR InterPro; IPR005540; KNOX1.
DR InterPro; IPR005541; KNOX2.
DR Pfam; PF03790; KNOX1; 1.
DR Pfam; PF03791; KNOX2; 1.
FT NON_TER 1
FT NON_TER 164
SQ SEQUENCE 164 AA; 17432 MW; 945B9D2936C9D45C CRC64;

Query Match 56.1%; Score 46; DB 2; Length 164;
Best Local Similarity 53.3%; Pred. No. 16;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 LKPPFKLKVEVFPFP 15
Db 19 LLPFPKVSQVYVTP 33

RESULT 15
Q948L7 PRELIMINARY; PRT; 164 AA.
ID Q948L7
AC Q948L7;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Transcription factor OSH3 (Fragment).
GN Name=OSH3;
OS Oryza rufipogon (Wild rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4529;
RN [1]
RP SEQUENCE FROM N.A.
RA Sato Y., Fukuda Y., Hirano H.;
```

RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AB071659; BAB68305.1; -.

DR Gramene; Q948L7; -.

DR GO; GO:0005634; C:nucleus; IEA.

DR GO; GO:0003677; F:DNA binding; IEA.

```
DR InterPro; IPR005540; KNOX1.
```

DR InterPro; IPR005541; KNOX2.

DR Pfam; PF03790; KNOX1; 1.

DR Pfam; PF03791; KNOX2; 1.

FT NON_TER 1

FT	NON-TER	164	164
FT	NON-TER	164	164

FILE	NON	LEA	104	104
SQ	SEQUENCE	164	AA;	17434 MW; 430B9D261998B6C5 CRC64;

Query Match
56.1%; Score 46; DB 2; Length 164;

Query Match 30.1%; Score 40, 25
Best Local Similarity 53.3%; Pred. No. 16;

BEST LOCAL
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
Similarity 55.3%; FREQ. NO. 18;

Qy 1 LKPFPLKVEVFPP 15

[illegible]

Search completed: April 12, 2005, 08:05:50
Job time : 73.1818 secs